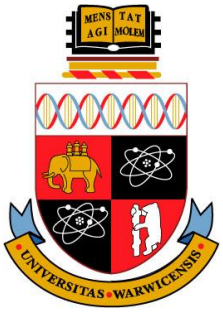


Biochemical Investigations of the Carotenoid Cleavage Dioxygenase Enzyme Family

Peter J. Harrison

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Life Sciences and Chemistry

University of Warwick, Department of Chemistry

THE UNIVERSITY OF WARWICK

October 2014

Supervisor: Professor T.D.H Bugg

I. Contents

II. List of Figures

III. List of Tables

IV. List of Equations

V. List of Abbreviations

AAo3	Abscisic acid aldehyde oxidase
ABA	Abscisic acid
ABA2	Xanthoxin dehydrogenase
ACO	Apocarotenoid cleavage oxygenase
APS	Ammonium persulfate
At	*Arabidopsis thaliana*
ATP	Adenosine triphosphate
AU	Arbitrary units
BCO	*β*-carotene oxygenase
BLAST	Basic local alignment search tool
BME	2-mercaptoethanol
CCD	Carotenoid cleavage dioxygenase
CCO	Carotenoid cleavage oxygenase
CRTISO	Prolycopene isomerase
Cys	Cysteine
D27	Dwarf27
Da	Dalton
° C	Degrees Celsius
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
DTT	Dithiothreitol
E	Enzyme
E. coli	*Escherichia coli*
EDC	1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ES	Enzyme-substrate complex

ESI	Electrospray ionisation
Ex	Excitation
Em	Emission
Ɛ	Extinction coefficient
FAD	Flavin adenine dinucleotide (oxidised)
FADH	Flavin adenine dinucleotide (reduced)
FPLC	Fast protein liquid chromatograph
GST	Glutathione-S-transferase
HA	Hydroxamic acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
HPLC	High pressure liquid chromatography
HPPD	*p*-Hydroxy-phenylpyruvate dioxygenase
HRMS	High resolution mass spectrometry
HRP	Horseradish peroxidase
hv	UV light
Hz	Hertz
IC_{50}	Half maximal inhibitory concentration
IPTG	Isopropyl *β*-D-1-thiogalactopyranoside
k_{cat}	Turnover number
K_{eq}	Equilibrium constant
K_M	Michaelis constant
kpsi	kilo pounds per square inch
LB	Luria broth
LC-MS	Liquid chromatography mass spectrometry
Le	*Solanum lycopersicum*
LSD	Lignostilbene dioxygenases
m/z	Mass to charge ratio

max	More axillary branching
MBP	Maltose binding protein
MES	2-(*N*-morpholino)ethanesulfonic acid
MS	Mass spectrometry
MOPS	3-(*N*-morpholino)propanesulfonic acid
N-His_6	Hexa-histidine
NAD	Nicotinamide adenine dinucleotide (oxidised)
NADH	Nicotinamide adenine dinucleotide (reduced)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NCED	9'-*cis*-epoxycarotenoid cleavage dioxygenases
NCEI	Neoxanthin isomerase
NDGA	Nordihydroguaiaretic acid
NEM	*N*-ethyl maleimide
NIC	Non-induced control
OD	Optical density
Os	*Oryza sativa*
Ox	Oxidised
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
pK_a	Acid dissociation constant
PMSF	Phenylmethanesulfonylfluoride
Red	Reduced
rms	Ramosous mutant
RNA	Ribonucleic acid
RPE65	Retinal pigment epithelium 65 kDa protein
S	Substrate
SDS	Sodium dodecyl sulfate

SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TBE	Tris-borate-EDTA
TBME	Tertbutylmethylether
TCEP	Tris-(2-carboxyethyl)phosphine
TEMED	Tetramethylethylenediamine
TP	Total protein
TSP	Total soluble protein
UV	Ultraviolet
UV-Vis	Ultraviolet-visible
Vu	*Vigna unguiculata*
λ	Wavelength
Zm	*Zea mays*

VI. Acknowledgements

First and foremost I would like to thank my supervisor, Prof. Tim Bugg, for all his help, enthusiasm and encouragement of the last three and-a-bit years. I would also like to thank Tim for all the opportunities that he has given me both during my PhD and during my undergraduate studies, for keeping me on track and reassuring me when things did not go to plan.

I would also like to offer my thanks to the many members of the Bugg group who have helped me over the years, both past and present. Special mentions must go to Dr Liz Hardiman for being an excellent role model and teaching me most of the things I know about molecular biology, Rahman RahmanPour for keeping me sane when everything was going wrong and listening to me moan, and Dr Paul Sainsbury for making the last three years well and truly enjoyable. Thanks also to everyone in the CBRF for their help, and especially to Anne Smith for keeping the place running!

Many people have help me directly with the project over the years, and none more so than my two awesome project students Flora and Sophie. Thank you both for all your hard work and all the chocolate you brought, it was a pleasure to have had the opportunity to work with you.

Thank you also to Dr Andrew Thompson and Jake Chandler at Cranfield University for all your help, suggestions and for providing the various enzymes in the project. The work done here would not have been possible without you both.

Thanks must also go to everyone at Syngenta working on the NCED project for supporting my work. A special thank you to Dr David Brocklehurst for all his help, suggestions and for making my visit to Jeallot's Hill enjoyable, and thank you to everyone in Biological Sciences at Jeallot's Hill for making me feel so welcome.

Completing this PhD would not have been possible without the support of so many people. Leanne and Ben, thank you so much for everything you have done for me over the last seven years, and especially for always being on hand with a glass of alcohol whether rain or shine. Your friendship has been more valuable than the words on this paper do justice. Thank you also to the rest of the Broomfield lot for the weekends of fun. Finally, and most importantly, I would like to thank family for all their love and support during over the years, especially my parents Ken and Alison for everything they have done. Oh, and my baby nephew Thomas. Your crying may drive me up the wall, but you have brought me so much happiness, thank you.

VII. Declaration

The work presented herein is my own work unless otherwise acknowledged and has not been submitted for a degree at another university.

Work in Chapter 3 on NCED-ABA2 coupled assays was completed partially at Syngenta Jealott's Hill in collaboration with Dr David Brocklehurst.

Work in Chapter 4 on the cloning of *Oryza sativa* D27 and initial assays was completed in collaboration with MChem project student Sophie Newgas.

Work in Chapter 5 on the establishment of CCD7 assays was completed in collaboration with Erasmus project student Flora Descombes.

Peter James Harrison

VIII. Abstract

The biosynthesis of the plant hormones strigolactone and abscisic acid is, in part, controlled by a family of enzymes known as the carotenoid cleavage dioxygenases (CCDs), which perform an oxidative cleavage reaction on a carotenoid substrate (9'-*cis*-neoxanthin for abscisic acid and 9-*cis*-*β*-carotene and 9-*cis*-*β*-apo-10'-carotenal for strigolactone) to form apocarotenoids, which are metabolised to the functional phytohormone. Phenotypic effects on seed dormancy and shoot branching have been observed in *Arabidopsis thaliana* and *Zea mays* on the application of a selection of hydroxamic acids based inhibitors, designed to inhibit CCDs, whereby application of the inhibitor to the plant result in a decrease in the time take for germination (abscisic acid mediated) or an increase in the number of lateral shoot branches (strigolactone mediated). However, the biochemical basis of these phenotypes is not understood.

In the present thesis, carotenoid cleavage dioxygenases from the abscisic acid biosynthesis pathway (9'-*cis*-epoxycarotenoid cleavage dioxygenases (NCED)) and strigolactone biosynthesis pathway (CCD7 and CCD8) were produced *in vitro* and assayed for inhibition against the hydroxamic acid inhibitors. The results show that *Z. mays* NCED is indeed inhibited by the hydroxamic acid inhibitors (D2: greater than 95% inhibition at 100 μM) in a time dependent fashion, indicating that inhibition of NCED is the basis of the seed germination phenotype. On the strigolactone biosynthesis pathway, recombinant *A. thaliana* CCD7 is not inhibited by the hydroxamic acids. However, recombinant *A. thaliana* CCD8 is inhibited by hydroxamic acids that show shoot branching phenotypes (D6 53% inhibition at 10 μM), suggesting that inhibition of CCD8 is the basis of the shoot branching phenotype. Structure activity relationships have also been performed to identify the key features of the hydroxamic acids required to inhibit each enzyme.

The biochemistry of several CCDs has also been investigated, along with that of the enzyme Dwarf27, an isomerase enzyme required for the isomerisation of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene on the strigolactone biosynthesis pathway. Investigations indicate that

Dwarf27 from *Oryza sativa* could be a novel iron-sulfur protein which isomerises *β*-carotene via a one electron transfer to or from the *β*-carotene substrate. D27 is also inhibited to some extent by certain hydroxamic acids (e.g. 41% by D30 at 100 μM). Biochemical characterisation of *A. thaliana* CCD8 provides evidence for a two-step mechanism involving acid-base catalysis, and evidence is obtained for an active site cysteine residue. A possible mechanism for the double oxidative cleavage reaction catalysed by CCD8 is proposed.

These results provide an insight into the biochemistry of the biosynthesis of the phytohormones abscisic acid and strigolactone and demonstrate that enzymes on the biosynthesis pathways of these hormones can be selectively inhibited using a chemical genetics approach. This has potential to aid the development of novel agrochemical compounds which could influence these processes to improve plant architecture and crop yield.

Chapter One: Introduction

1.1 Carotenoids

Carotenoids are C_{40} polyene terpenoids biosynthesised by phototrophic organisms such as plants, algae and certain bacteria (Walter & Strack 2011). Carotenoids are essential metabolites, so mammals must obtain carotenoids through their diet, mostly in the form of *β*-carotene (pro-vitamin A) and lycopene. To date, over 700 different carotenoids have been identified, which can be divided into two groups: Carotenoids which have an entirely hydrocarbon skeleton are termed carotenes, such as *β*-carotene (Figure 1.1), whilst those which are oxygenated are termed xanthophylls, such as neoxanthin (Figure 1.1). Xanthophylls which contain an epoxide group, such as neoxanthin, are termed epoxycarotenoids. Although structurally different, all carotenoids share the same polyene backbone.

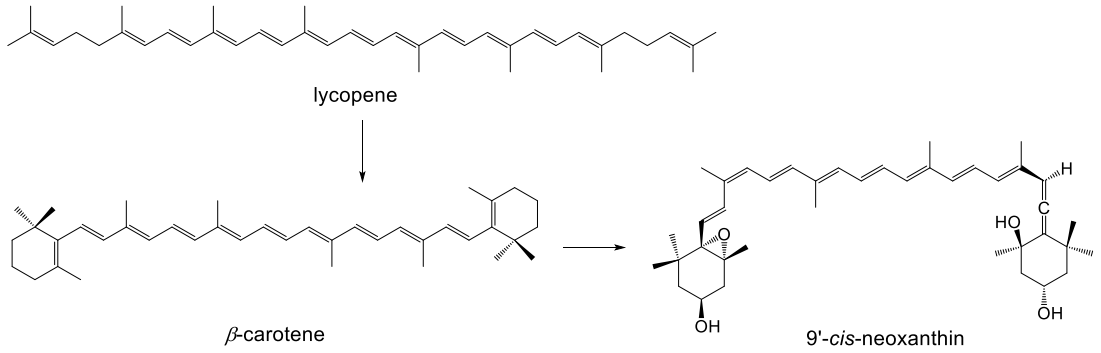


Figure 1.1 – Structures of the carotenoids lycopene and *β*-carotene and the epoxycarotenoid 9'-*cis*-neoxanthin.

The conjugated system of π electrons in carotenoids means that the energy of transition between the π and π* bonding and anti-bonding orbitals is low (Williams & Fleming 1995). As a result, carotenoids absorb in the high energy end of the UV-Vis spectrum (350-500nm), resulting in their yellow, red and orange colours. Carotenoids are found as accessory pigments in photosynthesis, where they can act as light harvesting pigments - whereby they absorb solar light energy and pass this energy onto chlorophyll by singlet-singlet excitation transfer (Cogdell & Frank 1987). In addition, carotenoids can quench triplet-chlorophyll, preventing the formation of the highly toxic singlet-oxygen ($^1O_2$) species. In a similar way,

carotenoids also act as antioxidants (Foote & Penny 1968). The highly conjugated nature of carotenoids results in a low lying π^* state, which renders the molecule susceptible to reactions with radicals. The resultant radical, which is resonance stabilised by the conjugation of the polyene backbone of the carotenoid, prevents the radical from causing oxidative damage (Burton 1989).

Carotenoids also have an ecological role as coloured pigments. The yellow and orange colours such as those produced by lutein and *β*-carotene respectively, are important in attracting pollinating insects to plants (Schaefer *et al*. 2004). Carotenoids have been reported to function in many other roles, such as reinforcing bacterial cell membranes and altering membrane properties (Rottem & Markowitz 1979, Britton *et al*. 2008). Due to the hydrophobic nature of carotenoids, it is not unexpected to find them localised to membranes.

1.2 Apocarotenoids

Carotenoids can be broken down via oxidative cleavage into smaller molecules known as apocarotenoids or norisoprenoids. The cleavage of carotenoids can occur either specifically or non-specifically. The non-specific cleavage of carotenoids occurs via either photo- or chemical oxidation. The specific cleavage of apocarotenoids, however, occurs via the action of a family of enzymes known as the carotenoid cleavage dioxygenases (CCDs). These apocarotenoids may be biologically active in their own right, or may be precursors to other biologically active molecules.

Apocarotenoids and their derivatives play a wide variety of different roles within biological systems. Possibly the most well studied apocarotenoid is retinal, one of the three forms of vitamin A (the other two being retinol and retinoic acid). Retinal, is produced by the symmetrical cleavage of *β*-carotene by the enzyme *β*-carotene oxygenase I (BCOI) and is the visual pigment in the rod cells of mammals, where it is bound via a Schiff base linkage to a lysine residue in a G-protein coupled receptor, called opsin. On exposure to light, 11-*cis*-retinilidene undergoes an 11-*cis* to all-*trans* isomerisation, which initiates the visual cascade,

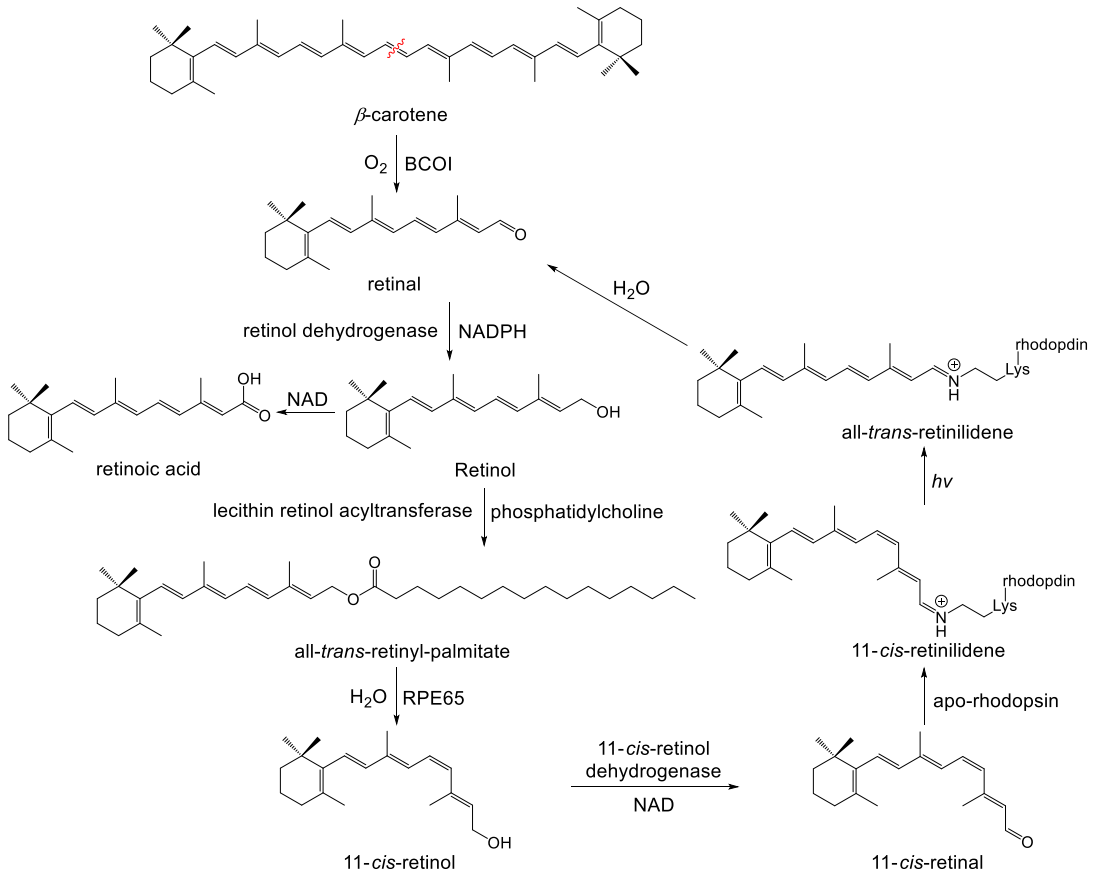


Figure 1.2 – Schematic representation of the visual pathway in mammals starting from *β*-carotene. The position of cleavage on *β*-carotene is indicated with a red line.

allowing vision (Figure 1.2) (Moiseyev *et al*. 2005, Jin *et al*. 2007, Wald 1968). Apocarotenoids also have roles as signalling molecules. For example, in plants, the signalling molecules strigolactone and abscisic acid are apocarotenoid derived molecules (Figure 1.3) (Umehara *et al*. 2008, Gomez-Roldan *et al*. 2008, Von-Lintig & Vogt 2000, Nambara & Marion-Poll 2005). Other biological roles of apocarotenoids have also been reported. For example, retinoic acid has been shown to have signalling properties in the immune system and in growth and development (Lampert *et al*. 2003, Spilianakis *et al*. 2005).

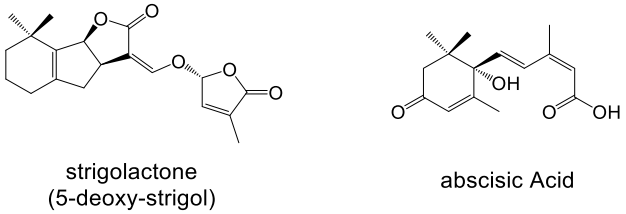


Figure 1.3 – Chemical structures of the phytohormones strigolactone and abscisic acid.

In many cases, apocarotenoids also have similar roles as carotenoids. Apocarotenoids are located in the thylakoid membrane of cyanobacteria and plants, acting as photosynthesis accessory pigments and as photoprotective compounds (Winterhalter & Straubinger 2000, Walter *et al*. 2010). Finally, examples of apocarotenoids acting as scent and aroma compounds have been reported, as have examples of apocarotenoids acting as colour pigments (Winterhalter & Straubinger 2000). Similarly, the volatile C_{13} apocarotenoid *β*–ionone is involved in the aroma of wine and tobacco (Walter *et al*. 2010, Winterhalter & Rouseff 2001).

1.3 Carotenoid Cleavage Dioxygenases

Carotenoid cleavage dioxygenases are a large family of non-heme iron (II) dependent enzymes which catalyse the oxidative cleavage of a double bond on the polyene backbone of carotenoids to produce two apocarotenoid products. Depending upon the position of cleavage on the backbone, this results in either aldehyde or ketone products being produced. As discussed, apocarotenoid and apocarotenoid derived molecules form important signalling molecules in plants and mammals. There is some debate within the literature as to whether CCDs are technically dioxygenases or monooxygenases, and as a result the names carotenoid cleavage dioxygenase and carotenoid cleavage oxygenase (CCO) are often used interchangeably. Here, however, the term CCD will be used.

Since the 1930s it has been known that *β*-carotene was the precursor to retinal. *In vivo* experiments in rats showed the production of retinal from *β*-carotene, which was proposed to occur through cleavage of the central 15,15' double bond. (Moore 1930, Karrer *et al*. 1930). Using cell extracts from rat liver and intestines, it was further shown in 1965 that this cleavage was the result of enzymatic activity of an enzyme termed *β*-carotene oxygenase (Goodman & Huang 1965, Olson & Hayaishi 1965). The gene encoding *β*-carotene oxygenase was not identified until the early 2000s, when it was shown to have sequence similarity with a plant gene termed vp14, which encodes the NCED1 enzyme from *Zea mays*. (von-Lintig & Vogt 2000, Wyss *et al*. 2000, Schwartz *et al*. 2007). Subsequently, large scale genomic sequencing has allowed the identification of large numbers of putative CCDs across all taxa (Auldridge *et*

al 2006b). However, in only a few cases are the biochemical functions of these enzymes understood.

Although they can exhibit substrate promiscuity, CCDs show high regio- and stereo-specificity for specific double bonds on carotenoid and apocarotenoid substrates. The regio-specificity of CCDs refers to the position on the carotenoid backbone where the CCD enzyme cuts. An enzyme with 9,10 cleavage specificity, for example, will cut between carbons 9 and 10 (or 9' and 10') (Figure 1.4).

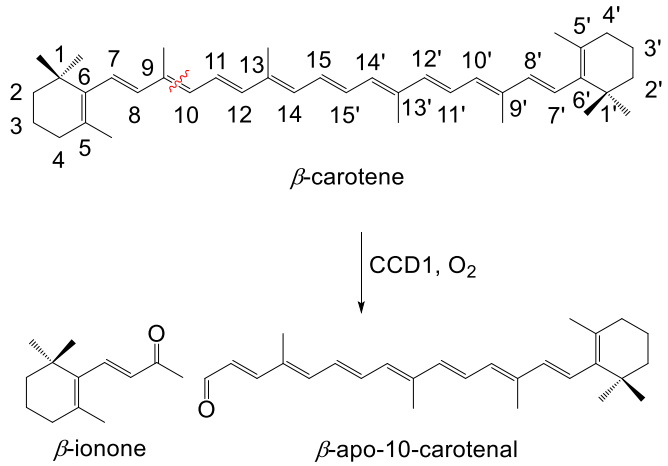


Figure 1.4 – Carbon numbering assignments for carotenoids, as represented on the structure of *β*-carotene in the cleavage reaction catalysed by CCD1. The position of cleavage on *β*-carotene is indicated with a red line.

To date, the majority of CCDs enzymes identified have been in plants, where they are involved in the biosynthesis of the signalling molecules strigolactone and abscisic acid. Within *Arabidopsis thaliana* (thale cress) nine distinct CCDs have been identified: CCD1, CCD4, CCD7, CCD8, NCED2, NCED3, NCED5, NCED6 and NCED9 (Neill *et al*. 1998, Tan *et al*. 2003, Sorefan *et al*. 2003, Booker *et al*. 2004). CCDs in other plant species are named according to homology to the *A. thaliana* enzymes. Only CCD1, 4, 7, 8 and NCEDs are found in all plants, which have varying numbers of NCEDs. For example, *A. thaliana* and *Z. mays* have five NCEDs, whereas *Solanum lycopersicum* (tomato) has only two. However, additional CCD enzymes may be found in certain plants. For example, zeaxanthin cleavage dioxygenase is found in *Crocus sativus* (crocus) and lycopene cleavage oxygenase is found in *Bixa orellana*

(bixa) (Bouvier *et al*. 2003). The plant CCD enzymes CCD1, CCD4, CCD7, CCD8 and NCEDs are discussed in more detail below.

In mammals, two CCD enzymes have been identified, termed *β*-carotene oxygenase I (BCOI) and *β*-carotene oxygenase II (BCOII) (Redmond *et al*. 2001, Yan *et al*. 2001, Kiefer *et al*. 2001). BCOI, identified in *Homo sapiens* and *Mus musculus*, is required for the 15,15' cleavage of *β*-carotene to produce retinal, which is required for vision. BCOII on the other hand, also identified in *H. sapiens* and *M. musculus*, has been shown to catalyse the 9,10 cleavage of *β*-carotene and lycopene. However, the biological role of BCOII remains unknown. A third protein with sequence similarity to the CCD family is also present in mammals. Retinal pigment epithelium 65 kDa protein (RPE65) is an iron dependent enzyme which is structurally related to the CCDs (Moiseyev *et al*. 2006). However, RPE65 does not catalyse the cleavage of carotenoid or apocarotenoids. Instead, RPE65 is required for the isomerisation of all-*trans*-retinyl ester to 11-*cis*-retinol in the visual cycle (Figure 1.2). RPE65 is discussed in more detail in Section 1.7.

NinaB is a CCD enzyme identified from *Galleria mellonella* (great wax moth) and is similar to BCOI in that is required for the biosynthesis of retinal (Oberhauser *et al*. 2008). However, NinaB is interesting in that it also appears to catalyse the isomerisation of the apocarotenoid products (Figure 1.5). NinaB has been shown to be able to catalyse the cleavage

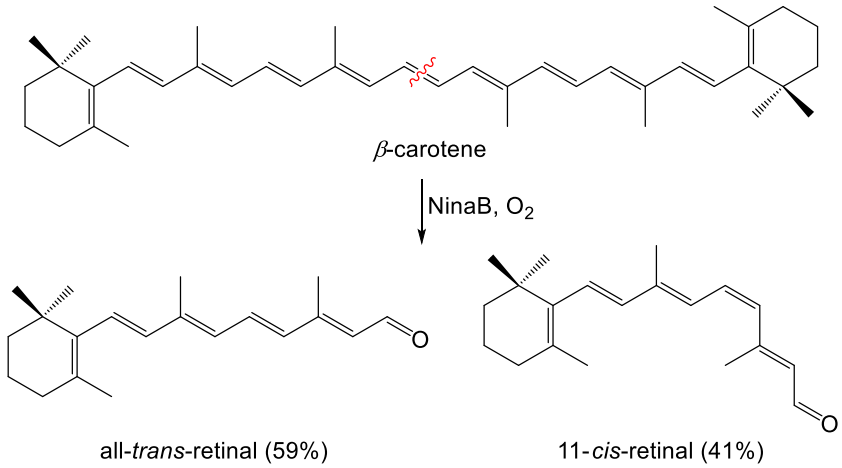


Figure 1.5 – Isomerisation and cleavage reaction of *β*-carotene catalysed by *G. mellonella* NinaB. The all-*trans*-retinal isomer is formed in preference to the 11-*cis* isomer. The position of cleavage on *β*-carotene is indicated with a red line.

of a range of carotenes and epoxycarotenoids structurally related to *β*-carotene to all-*trans* and 11-*cis* retinal.

Many CCD enzymes have been identified in bacteria. One of the first bacterial CCDs to be investigated was the cyanobacterial apocarotenoid cleavage oxygenase (ACO) from *Synechocystis sp*. PCC 6803 (Ruch *et al*. 2005). ACO is involved in the biosynthesis of retinal from the substrate all-*trans*-*β*-apo-8'-carotenal via a 15,15' cleavage reaction (Figure 1.6). *In vitro* experiments show that ACO is selective for all-*trans* substrates and will accept oxygenated derivatives of the *β*-apo-8'-carotenal substrate, but will not cleave either *β*-carotene or lycopene. Changes in the chain length of the substrate from that of the natural *β*-apo-8-carotenal alter the rate of the cleavage reaction. If the length of the substrate was either increased or decreased then the reaction rate falls.

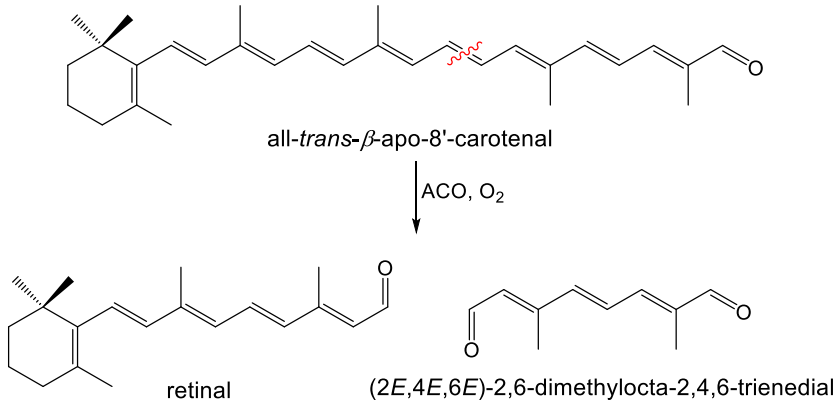


Figure 1.6 – Cleavage reaction of all-*trans*-*β*-apo-8'-carotenal to retinal and (2*E*, 4*E*, 6*E*)-2,6-dimethylocta-2,4,6-trienedial. The position of cleavage on all-*trans*-*β*-apo-8'-carotenal is indicated with a red line.

Other bacterial CCDs have also been identified in *Mycobacterium tuberculosis* and *Novosphingobium aromaticivorans* (Scherzinger *et al*. 2010, Kim *et al*. 2012). The *M. tuberculosis* CCD is somewhat interesting in that it only cleaves carotenoids with terminal cyclic groups at both the 15,15' and 13,14 positions (Kim *et al*. 2012). This regio-promiscuity is somewhat unusual for the CCD family. It could be reasoned that the terminal cycle group is required for binding, and that the active site architecture is such that it provides only a loose fit of the substrate within the active site. A fungal CCD, termed CarX has also been identified in *Gibberella fujikuroi* (Prado-Cabrero *et al*. 2007).

1.4 Structural Aspects of CCDs

To date, two CCD enzymes have been crystallised. These structures belong to *Synechocystis sp.* PCC 6803 ACO (PDB: 2B1W) and to *Z. mays* VP14 (PDB: 3NPE) (Figure 1.7) (Kloer *et al.* 2005, Messing *et al.* 2010). A third structure, belonging to *Bos taurus* RPE65 has also been reported (PDB: 3FSN) (Kiser *et al.* 2009). Although RPE65 does not catalyse the cleavage of carotenoids, it shares both sequence and structural similarity with the CCD family. Analysis of the structures of ACO, VP14 and RPE65 reveals that CCDs have a

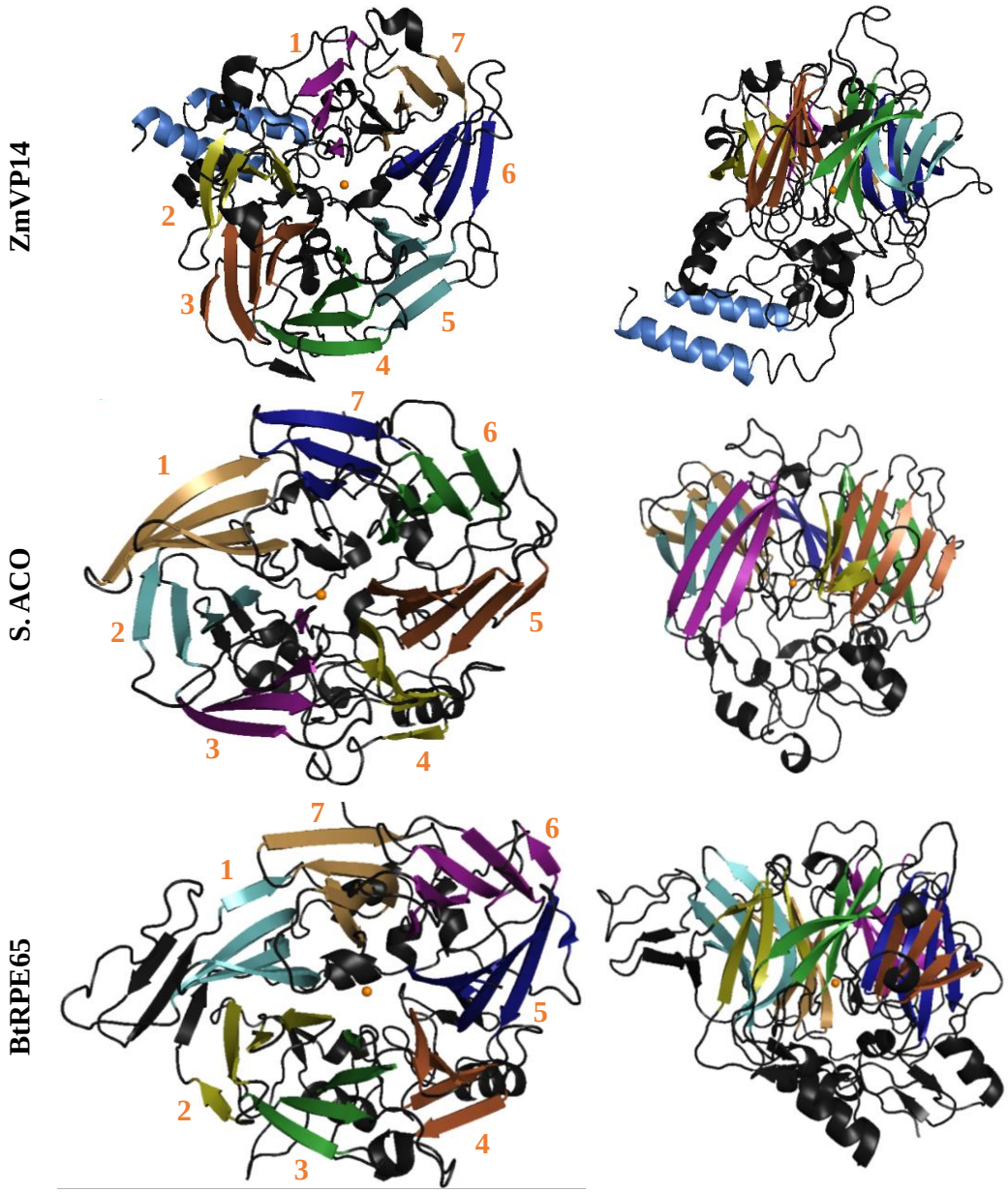


Figure 1.7 – Structures of the CCD enzymes *Zea mays* VP14 (PDB: 3NPE) and *Synechocystis sp.* ACO (PDB: 2B1W), and the retinyl isomerase *Bos taurus* RPE65 (PDB: 3FSN). *β*-sheet blades are numbered in orange.

conserved 7 bladed β propeller motif (Figure 1.7). The blades of this β-propeller consist of 4 to 5 β sheets arranged in an antiparallel fashion. At the base of the β-propeller sits the catalytic iron centre, buried deep within the enzyme, most likely to prevent escape of radical species which could damage the cell.

The iron (II) centre is chelated by four histidine residues from the innermost blades of the β propellers in an octahedral fashion (Figure 1.8). Surrounding this are second shell glutamate or aspartate residues which are partially conserved. The tetradentate co-ordination state is unusual for non-heme iron enzymes, which usually have tridentate coordination by

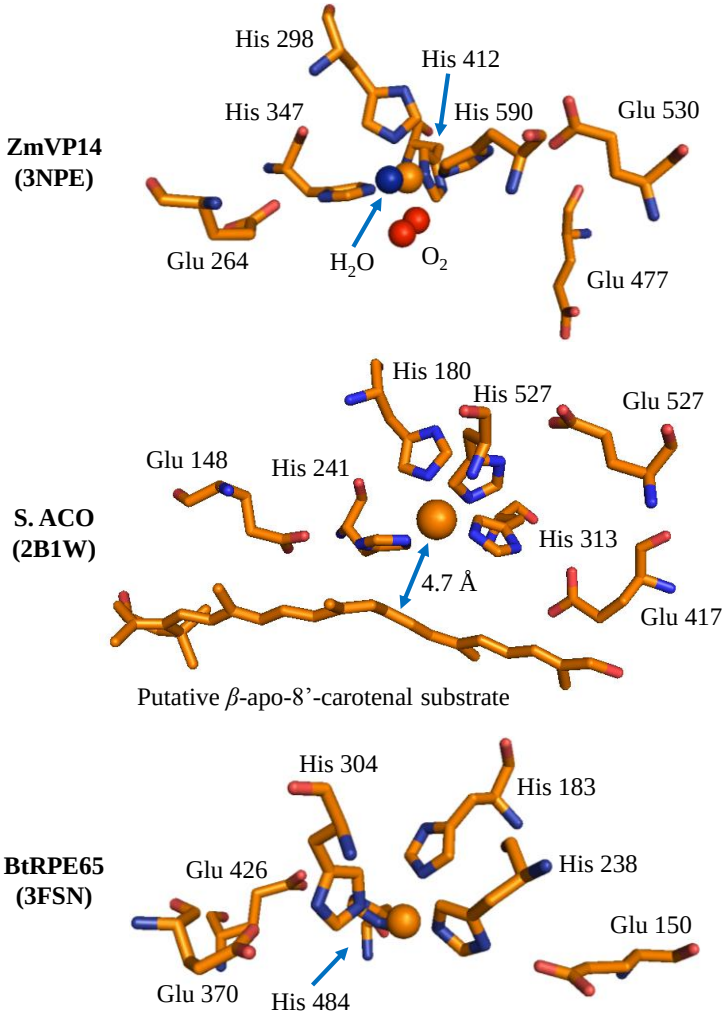


Figure 1.8 – Active site structures of the CCD enzymes *Z. mays* VP14 and *Synechocystis sp*. ACO and the retinyl isomerase *Bos taurus* RPE65. The four histidine residues coordinating the iron are strictly conserved, whilst the second shell residues are semi conserved. In VP14, water and oxygen have been modelled into the active site (not observed crystallographically). The putative β-apo-8'-carotenal substrate is also shown in the *Synechocystis* structure, positioned 4.7 Å below the iron.

histidine, glutamate or aspartate residues. Where there is tetradentate coordination, such as in the case of the intradiol protochatechuate 3,4-dioxygenase from *Pseudomonas putida*, the coordination is performed by a variety of ligands (Ohlendorf *et al.* 1988). For the *P. putida* protochatechuate dioxygenase this is by two histidine residues and two tyrosine residues. Only a handful of enzymes are known to have tetradentate coordination by four histidine residues, such as superoxide reductase (Gilmor *et al*. 1997). Sequence alignment of the CCD enzyme

```
BtRPE65 (Q28175) CTETNF TAHPH SYVHSFG NLFHHIN NTYEDHE YPSEPIF VTFHGLF
HsRPE65 (Q16518) CTETNF TAHPH SYVHSFG NLFHHIN NTYEDNG YPSEPIF VTFHGLF
HsBCO2  (Q9BYV7) CTETNF TAHPH SYYHSFG VTFHQIN NAFEDQG YPSEPVF YGFHGTF
HsBCO1  (Q9HAY6) TSETNY TSHPH SYYHSFG VVFHHVN NAYEEDG WPAEPLF MDLHGLF
ZmNCED1 (O24592) MSEDDL IAHPK TMIHDFA FCFHLWN NAWEDEA FGGEPCF FGFHGTF
ZmCCD1  (Q45VT7) LSEADK TAHPK VMMHDFA FIFHNAN NAWEEGD FGSEAIF YGFHAFF
AtNCED9 (Q9M9F5) MSEDDL IAHPK TMIHDFA FCFHLWN NSWEEPE YGGEPLF YGFHGTF
AtNCED3 (Q9LRR7) MSEDDL IAHPK TMMHDFA FCFHLWN NAWEEPE YGGEPLF YGFHGTF
AtNCED5 (Q9C6Z1) MSEDDL IAHPK TMMHDFA FCFHLWN NAWESPE YGGEPFF YGFHGTF
AtNCED2 (O49505) MSEDDL IAHPK TMIHDFA FCFHLWN NAWESPE YGGEPFF YGFHGTF
AtNCED6 (Q9LRM7) MSEDDL IAHPK TMIHDFA FCFHLWN NAWEERT FGGEPCF YGFHGTF
AtCCD4  (O49675) LGESDL TAHPK SFLHDFA NIIHAIN NAWDEDD YGGEPFF YGFHGLF
AtCCD1  (O65572) LQEADK TAHPK IMMHDFA FIFHNAN NAWEEED YGSEAIY YGFHALF
AtCCD7  (Q7XJM2) LWEGGE LSHYK MMIHDWA WLIHSGN NAYETRE FVGEPMF ICVHSFY
AtCCD8  (Q8VY26) LTETQK SAHPI GWVHSFA VTFHFIN NAYEEDK IPSEPFF YGLHGCW
S.ACO   (P74334) LWEGGQ SAHPR AFIHDFA FVFHHAN NAFEEN- FAGEPIF YPLHGSW
DmNinaB (Q9NKW9) FTETPF TSHPH GYMHTFG FYLHIIN NCFERDG YPSEPIF KCLHGWF
LeCCD1a (Q6E4P5) LSEADK TAHPK IMMHDFA FIFHNAN NAWEEGD FGSEAVF YGFHAFF
PsNCED2 (Q8LP16) MSEDDI IAHPK TMMHDFA FCFHLWN NAWEEPE FGGEPLF YGFHGTF
PsNCED3 (Q8LP15) MSEDDL IAHPK TMIHDFA FCFHLWN NAWEEPE FGGEPFF YGFHGTF
PaNCED1 (Q9AXZ3) MSEDDL IAHPK TMIHDFV FCFHLWT TAWEEPE YGGEPYF YGFHGTF
PaNCED3 (Q9AXZ4) MSEDDL IAHPK TMIHDFA FCFHLWN NAWEEPE FGGEPYF YGFHGTF
LeNCED2 (C5H805) MSEDDV IAHPK TMMHDFA FCFHLWN NAWEEPE YGGEPLF YGFHGTF
BoLCO   (Q7OYP8) LCEYDL TAHPK SFVHDFA NMVHVVN NAWEEEG FGGEPFF YGFHGLF
CsZCD   (Q84K96) LCEYDL TAHPK SFVHDFA NMVHVVN NAWEEEG FGGEPFF YGFHGLF
PsCCD1  (Q8LP17) LSEGDK TAHPK VMMHDFA FIFHNAN NAWEEED FGSEAVY YGFHAFF
MtCCO   (O06785) LVEAGV TAHPQ PMMHSFS YVYHPLN NAYSECR LIGEMVF MGFHGNW
MmBCO2  (Q99NF1) STETNF TAHPH SYYHSFG LTYHQIN NAFEDQG YPSEPVF YGFHGTF
MmBCO1  (Q9JJS6) TTETNY TSHPH SYYHSFG VVFHHVN NAYEEDG WPAEPLF LDLHGLF
NaCCO   (Q2GA76) MKEDSP TAHPK CMMHDFG FASHVLN NAWQEGT SLQEPCF FGLHGNW
PpLSD   (Q53353) MKEDSP CAHPK CMMHDFG FVGHVMN NAFNDGT AIQEPCF QGLHGNW
LeNCED1 (O24023) MSEDDL IAHPK TMMHDFA FCFHLWN NAWEEAE YGGEPLF YGFHGTF
GfCarX  (Q5GN50) TCESGP TGHPK RMMHDFG CIFHTAN NTWDSQS YAQEPRF YGLHGTW
```

Figure 1.9 – Partial T-Coffee (Notredame *et al*. 2000) sequence alignment of carotenoid cleavage dioxygenases showing conserved histidine residues (orange) and semi-conserved second shell residues (red). Conflicts from the consensus are in blue. Accession numbers are shown in brackets. Bt – *Bos taurus*; Hs – *Homo sapiens*; Zm - *Zea mays*; At – *Arabidopsis thaliana*; S. – *Synechocystis sp.*; Dm – *Drosophilia melanogaster*; Le – *Solanum lycopersicum*; Ps – *Pisum sativum*; Pa – *Persea americana*; Bo – *Bixa orellana*; Cs – *Crocus sativum*; Mt – *Mycobacterium tuberculosis*; Mm – *Mus musculus*; Na – *Novosphingobium aromaticivorans*; Pp – *Pseudomonas paucimobilis*; Gf - *Gibberella fujikuroi*. A full sequence alignment is provided in appendix 10.1

family shows that the four histidine residues required for iron coordination are strictly conserved across the family (Figure 1.9). Since tetradentate coordination leaves only two vacant sites for catalysis, it is reasonable to assume that the tetradentate coordination is essential for the CCD enzyme family. The crystal structure of ACO in the absence of iron shows that the histidine ligands and second shell residues remain in the same position, suggesting a rigid structure. The second shell residues may either participate in catalysis or modulate the redox potential of the iron.

Within the structures of CCD enzyme there are two tunnels. The first tunnel runs from the surface of the enzyme to the active site (Figure 1.10). This tunnel is lined with polar residues and is believed to be required for delivering oxygen to the active site. The second tunnel is a long hydrophobic tunnel which runs perpendicular to the axis of the β-propeller

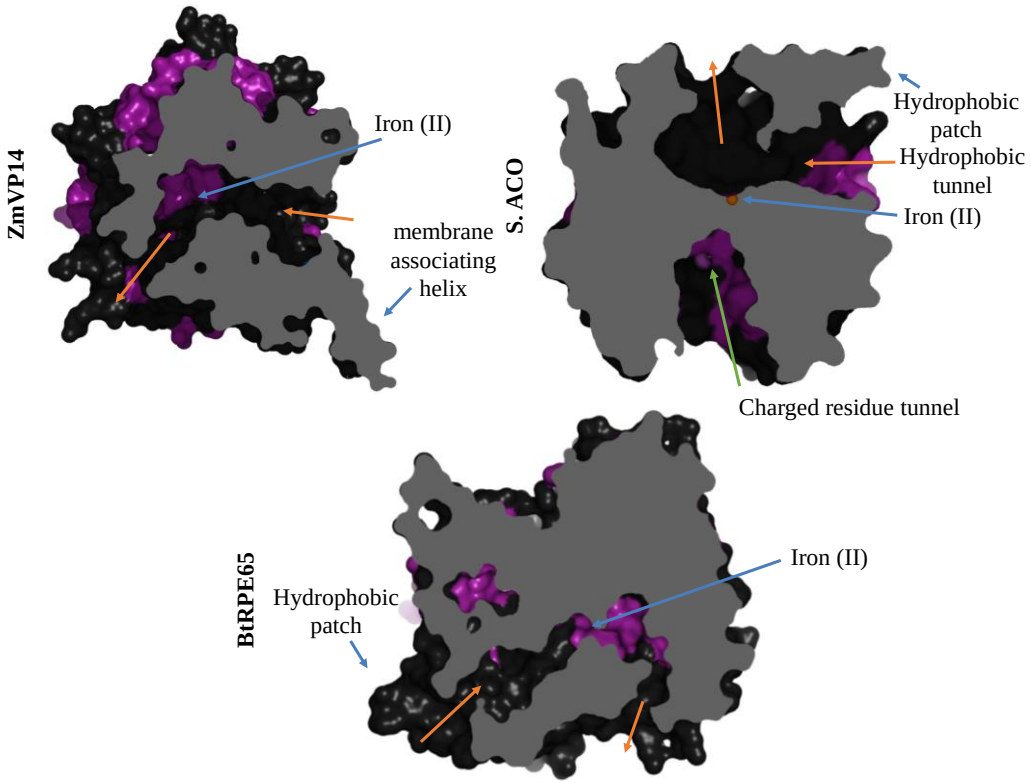


Figure 1.10 – Cross-section of the CCD enzymes *Zea mays* VP14 and *Synechocystis sp*. ACO and the retinyl isomerase *Bos taurus* RPE65 showing the substrate binding tunnels. Tunnel entrances and exits are shown with orange arrows. The location of the catalytic iron centre is shown with blue arrows (iron is only visible in ACO due to orientation of structures) as an orange sphere. Tunnel entrances are located next to hydrophobic regions, allowing substrate extraction from the membrane. Purple patches represent β-sheet blades.

from the surface of the enzyme to the catalytic iron at the centre (Figure 1.10). At the iron centre, the tunnel turns 90° and exits the enzyme parallel to the β-propeller. This is the substrate binding tunnel and is required to deliver the carotenoid substrate to the active site. Analysis of the VP14 structure shows that the carotenoid substrate is held in place by non-specific hydrophobic van der Waals interactions with leucine, valine and isoleucine residues within the binding tunnel (Messing *et al.* 2010). Additionally, the cleavage site is surrounded by three phenylalanine residues. These phenylalanine residues, along with additional non-specific hydrophobic interactions from methionine, leucine, valine, tryptophan and proline residues aid in the binding of the substrate. Mutation of the analogous three phenylalanine residues which surround the cleavage site in the structurally related *Z. mays* CCD1 enzyme to alanine residues results in a decrease in activity by up to 30%, suggesting that the residues are indeed important for substrate binding. From modelling with the ACO structure, it appears that the length of the tunnel before it turns to exit the enzyme and the size of the opening to the tunnel on the surface of the enzyme helps to dictate substrate specificity (Kloer *et al*. 2005). A small opening blocks the bulky terminal cyclic groups on carotenoids from entering the active site. This essentially turns the tunnel into a ruler, with the opening acting as a stopper, ensuring the correct double bond is positioned for cleavage.

In their original 2005 paper on the structure of ACO, Kloer *et al*. crystallised ACO in the presence of apocarotenoid substrates, and assigned electron density observed within the active site of the enzyme to a carotenoid substrate bound within the active site (Kloer *et al.* 2005). Interestingly, this electron density did not give a straight linear chain as would be expected for a carotenoid (Sui *et al*. 2013). Instead, it was proposed that the carotenoid had undergone isomerisation upon binding to the enzyme. As a result, analysis of the structure produced by Kloer *et al*. suggested that the bound substrate is positioned approximately 5.0 Å from the iron (Figure 1.8). This would imply that the interaction between the substrate and the activated oxygen species would occur distal to the iron (II) centre, with the activated oxygen species sitting between the iron and substrate. However, Sui *et al*., crystallising ACO in the

presence of a wide variety of different surfactants and in the absence of carotenoids or apocarotenoids, have observed the same electron density within the active site (Sui *et al*. 2014). Consequently the electron density within the active site may in fact be due to either detergent or ordered water molecules sitting within the hydrophobic active site. Hence the geometry of the carotenoid substrate within the active site is still uncertain.

Within the VP14 crystal structure, the two remaining coordination sites-on the iron centre are proposed to be occupied by water and dioxygen, which is bound in a side on fashion (Messing *et al*. 2010). In the ACO crystal structure no other ligands were reported on the iron other than the four histidine residues. Given the observations from the VP14 structure, it is reasonable to hypothesise that the oxygen binds to the iron in a side on fashion, sitting between the iron and the carotenoid substrate.

Structural analysis also reveals that both the Vp14 and ACO enzymes are targeted to the membrane. In Vp14, an α-helical domain, consisting of two long antiparallel α-helices lined with non-polar alanine, phenylalanine and leucine residues, targets the enzyme to the active site (Figure 1.11). Given the hydrophobicity of the 9'-*cis*-neoxanthin substrate, which is likely to reside within the plastid membrane, this would allow the enzyme easy access to

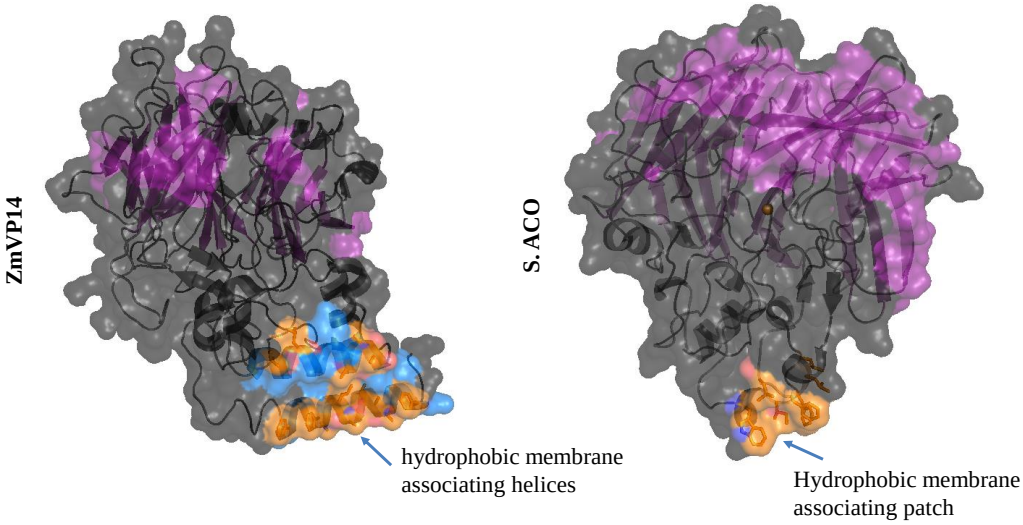


Figure 1.11 – Crystal structures of the CCD enzymes *Z. mays* VP14 and *Synechocystis sp*. ACO showing hydrophobic patches required for membrane association and access to carotenoid and apocarotenoid substrates.

the substrate, whilst preventing the substrate from entering the polar cytosol. A similar hydrophobic path is found on the surface of ACO (Figure 1.11). Here the patch is lined with leucine and phenylalanine residues, which are predicted to allow the enzyme to dip into the membrane and access substrate. Given that ACO acts on apocarotenoids, which are more polar due to the presence of either an aldehyde or ketone from a subsequent cleavage, it seems reasonable that ACO can access either substrates from either the cytosol or the membrane (Kloer *et al.* 2005, Messing *et al.* 2010).

Sequence alignment of CCDs shows that they are related to another family of enzymes known as lignostilbene dioxygenases (LSD) (Figure 1.9) (Kamoda & Saburi 1993). LSD enzymes are found in bacteria and are required for the cleavage of lignostilbene into two equivalents of vanillin, which may be involved in the breakdown of lignin (Figure 1.12) (Bugg *et al*. 2011). Sequence alignment shows that LSD contains the four conserved histidine residues and second shell glutamine residues which are found in the CCD family (Tan *et al*. 1997). LSD is known to be an iron dependent enzyme and given the sequence similarity between the enzymes it is possible that they share a common ancestor and catalytic mechanism.

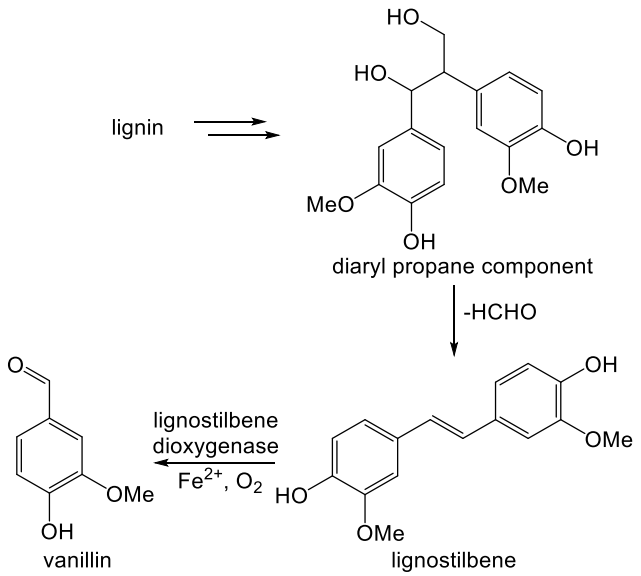


Figure 1.12 – Representation of the breakdown of lignostilbene to vanillin, which may be involved in the lignin degradation pathway performed by bacteria and fungi.

1.5 Plant Carotenoid Cleavage Dioxygenases: Biochemical Roles

1.5.1 Carotenoid Cleavage Dioxygenase 1 (CCD1)

The carotenoid cleavage dioxygenase 1 (CCD1) enzyme subfamily is unusual amongst the plant CCDs, as it is the only CCD not to be localised to the plastids within the plant cell (Bouvier *et al*. 2005). Originally identified in *A. thaliana*, AtCCD1 is a cytosolic enzyme, which catalyses the cleavage of the 9,10 or 9',10' bond in *β*-carotene to give a C_{13} ketone (*β*-ionone) and a C_{27} aldehyde (Figure 1.13) (Schwartz *et al*. 2001). Other CCD1 enzymes have been identified, based on sequence similarity to AtCCD1 (Schwartz *et al.* 1997). CCD1 enzymes have been identified in *Cucumis melo* (melon), *S. lycopersicum*, *Oryza sativa* (rice), *Coffea arabica* (coffee), *Vitis vinifera* (grape), *C. medica* (citrus fruit), *Fragaria ananassa* (strawberry) and *Rosa damascena* (rose) (Idbah *et al*. 2006, Simkin *et al*. 2004, Ilg *et al*. 2009, Simkin *et al*. 2009, Mathieu *et al*. 2005, Kato *et al*. 2006, Garcia-Limones *et al*. 2008, Huang *et al*. 2009).

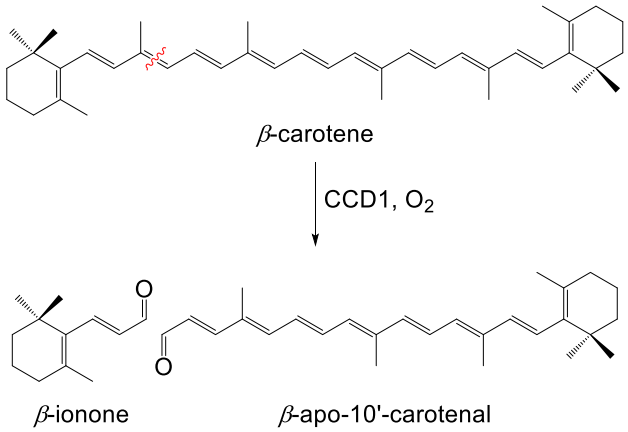


Figure 1.13 – Cleavage reaction of *β*-carotene to *β*-ionone and *β*-apo-10'-carotenal catalysed by CCD1. The position of cleavage on *β*-carotene is indicated with a red line.

In *A. thaliana*, CCD1 has been shown to be expressed in high levels in flowers and siliques (Auldridge *et al*. 2006a). This supports the idea that one of the roles of CCD1 is the biosynthesis of aroma volatiles such as *β*-ionone, required for attracting pollinating insects. The expression of *A. thaliana* CCD1 has also been shown to be brought about by drought stress, suggesting that CCD1 may have a role in the plant drought tolerance response (Neill *et al*. 1998).

Interestingly, all members of the CCD1 enzyme family identified to date have a high degree of sequence similarity, indicating shared substrate specificities (Tan *et al*. 1997, Idbah *et al*. 2006). The CCD1 enzyme family shows a high degree of substrate promiscuity, with 9,10 cleavage activity having also been reported against numerous carotenoids including, lycopene, 9'-*cis* and all-*trans*-violoxanthin, 9'-*cis*-neoxanthin and apocarotenoids including all-*trans*-*β*-apo-8'-carotenal (Figure 1.14) (Schwartz *et al.* 2001, Auldridge *et al.* 2006a, Schmidt *et al*. 2006). In addition to 9,10 cleavage, 5,6 and 5',6' cleavage has been reported *in vitro* on the substrate lycopene. However, this has not been shown *in vivo* (Vogel *et al*. 2008).

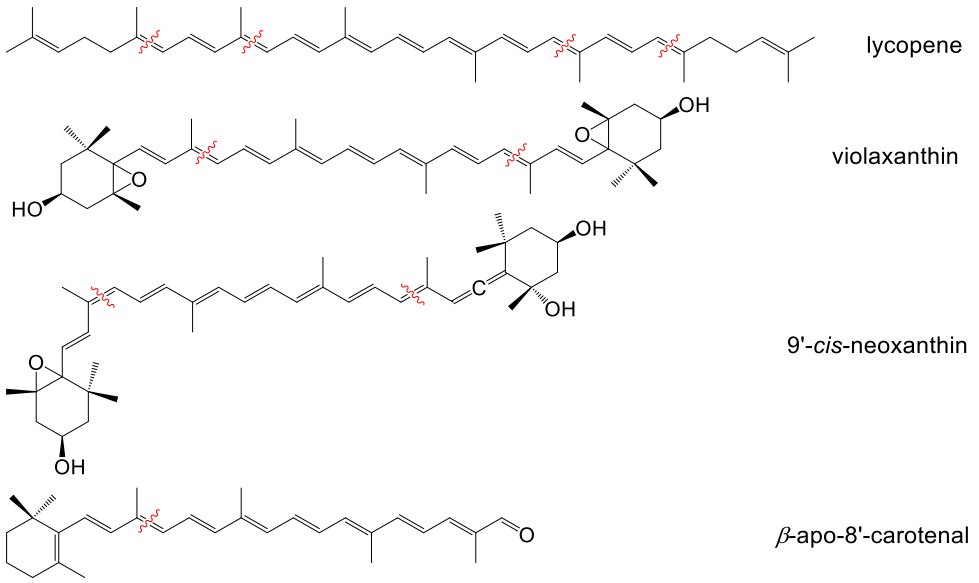
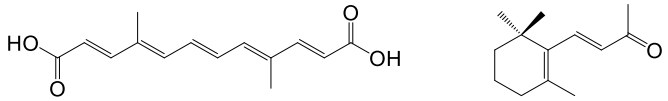


Figure 1.14 – Cleavage positions of CCD1 on the carotenoid substrates lycopene, violaxanthin, 9'-*cis*-neoxanthin and the apocarotenoid *β*-apo-8'-carotenal.

Deletion of the gene encoding CCD1 in *A. thaliana* does not result in a change in plant phenotype, indicating that in plants it is likely that CCD1 is not the sole enzyme responsible for the cleavage of *β*-carotene (Auldridge *et al*. 2006a). Suppression of CCD1 by RNA interference (RNAi) *in planta* results in only a 50% loss of C_{13} (*β*-ionone) apocarotenoid levels (Floss *et al*. 2008). In addition, the CCD7 enzyme has also been shown to catalyse a 9,10/9',10' cleavage of *β*-carotene to give *β*-ionone, and a C_{27} apocarotenoid (Schwartz *et al*. 2004). It is likely, due to the hydrophobicity of the *β*-carotene substrate, that CCD7 predominantly catalyses the cleavage of *β*-carotene in or around the plastid membranes. CCD1 is then able to act on the more polar C_{27} apocarotenoid in the cytosol, catalysing a 9,10

cleavage to give a second equivalent of *β*-ionone and a C_{14} dialdehyde. CCD1 is important for the biosynthesis of C_{13} and C_{14} apocarotenoids in the roots, such as *β*-ionone and mycorradicin, which form important plant pigments and volatiles (Bouvier *et al*. 2005) (Figure 1.15). The possibility also exists that CCD1 acts on other C_{27} apocarotenoids, derived from carotenoids other than *β*-carotene, leading to other apocarotenoids.



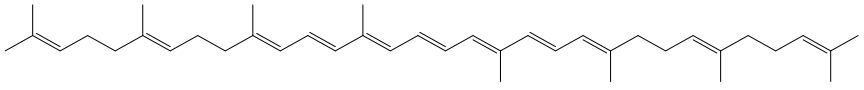
mycorradicin *β*-ionone

Figure 1.15 – Structures of the aroma volatiles mycorradicin and *β*-ionone.

1.5.2 Carotenoid Cleavage Dioxygenase 4 (CCD4)

The carotenoid cleavage dioxygenase 4 family of CCDs is possibly the least explored member of the CCD family. Little is known about the *in planta* roles of CCD4 or its biochemistry. RNAi silencing of the CCD4 gene in *Chysanthemum morifolium* (chrysanthemum) resulted in no phenotypic changes other than decolourisation of petals from white to yellow (Ohmiya *et al*. 2006). Further *in planta* experiments using RNAi in *Solanum tuberosum* (potato) resulted in increased levels of violoxanthin, indicating that violoxanthin may be the natural substrate *in vivo* (Campbell *et al*. 2010). Phenotypic effects were observed on heat sprouting and tubers, suggesting that the CCD4 cleavage product may be involved in development of potato tubers.

In vivo experiments with *β*-carotene, *cis*-*ζ*-carotene (Figure 1.16) and lycopene accumulating *Escherichia coli* revealed that CCD4 from *Malus domestica* (apple), *C. morifolium*, *R. damascena*, *Osmanthus fragrens* (osmanthus) and *A. thaliana* were only able to cleave *β*-carotene at the 9,10 / 9',10' position, producing *β*-ionone and *β*-apo-10-carotenal (Huang *et al*. 2009, Rubio *et al*. 2008). *In vitro*, CCD4 enzymes from *A. thaliana* and *R. damascena* were both able to cleave *β*-apo-8'-carotenal at the 9,10 / 9',10' position to produce



ζ–carotene

Figure 1.16 – Structure of *ζ*-carotene.

β-ionone. Given the difference in *in vitro* specificity, it is possible the different CCD orthologs from different species act on different substrates *in planta*.

1.5.3 Carotenoid Cleavage Dioxygenase 7 and 8 – Strigolactone Biosynthesis

CCD7 and CCD8 are often discussed together since both are involved in the biosynthesis of the plant signalling molecule strigolactone. The existence of strigolactone has been known for some time, as it is involved in rhizosphere signalling in plants. However, it was not until recently that strigolactone was identified as a phytohormone which acts to inhibit lateral shoot branching in plants (Akiyama *et al*. 2005, Gomez-Roldan *et al*. 2008, Umehara *et al*. 2008, Brewer *et al*. 2013). Currently, only the details of the initial stages of strigolactone biosynthesis are known.

Analysis of More Axillary Branching (*max*, i.e. more shoot branching) mutants and Ramosous mutants (*rms*) of *A. thaliana* and *Pisum sativum* (garden pea) respectively led to the identification of *max*/*rms* genes required for the biosynthesis of a then-unknown hormone (Turnbull *et al*. 2002, Beveridge 2000, Beveridge *et al*. 1996, Morris *et al*. 2001). The product of *max*3 and *max*4 was found to be graft transmissible, indicating that the product of the *max* pathway was a signalling molecule (Leyser 2005, Dun *et al*. 2006). Initially *max*4/*rms*1 was identified as a carotenoid cleavage dioxygenase (CCD8), however, its biochemical function remained unknown (Sorefan *et al*. 2003). *In vivo* experiments by Auldridge *et al*. suggested CCD8 was capable of cleaving the carotenes lycopene, zeaxanthin (Figure 1.17) and *β*-carotene (Auldridge *et al*. 2006). Subsequently *max*3/*rms*2 was shown to code for a carotenoid cleavage dioxygenase, termed CCD7, and was shown to cleave *β*- and ζ-carotene *in vivo* (Booker *et al*. 2004). *In vitro* biochemical characterisation by Schwartz *et al*. on the gene

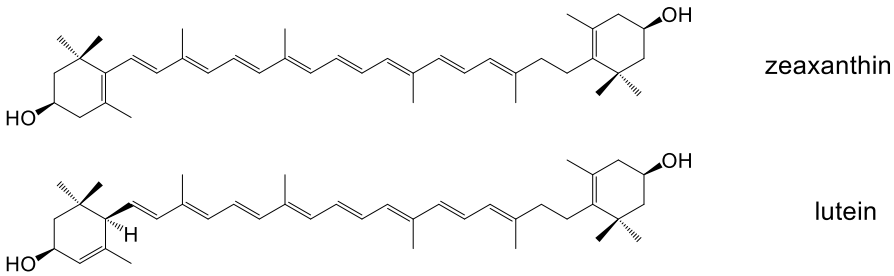


Figure 1.17 – Structures of zeaxanthin and lutein.

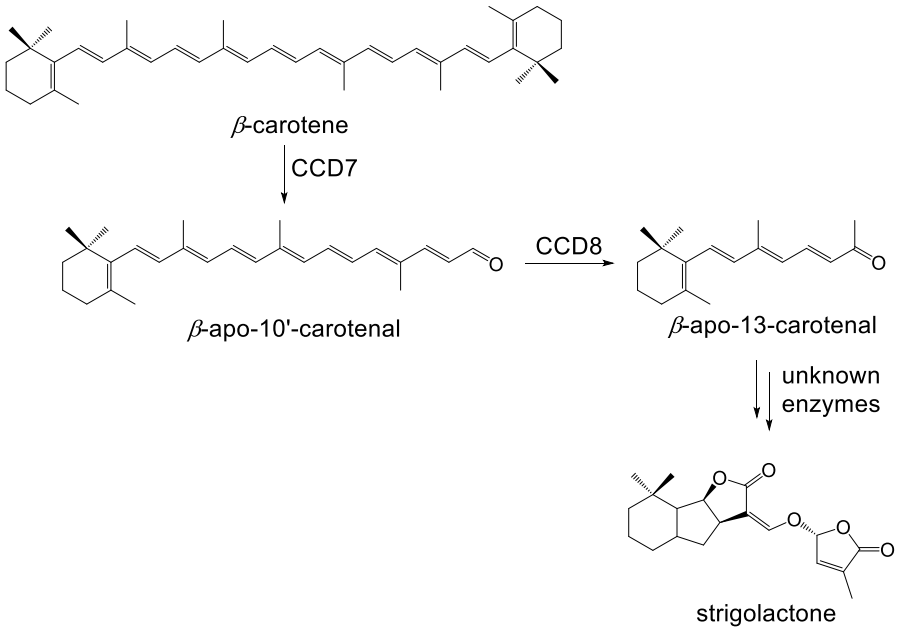


Figure 1.18 – Initial proposal for the start of the biosynthetic route to strigolactone, via the formation of *β*-apo-13-carotenal produced by the sequential action of CCD7 and CCD8.

products of *max*3/*max*4 (*A. thaliana* CCD7 and CCD8 respectively) showed that CCD7 was indeed able to cleave all-*trans*-*β*-carotene at the 9,10/9',10' position to produce all-*trans*-*β*-apo-10-carotenal (a C_{27} aldehyde) and that CCD8 then cleaved *β*-apo-10-carotenal to produce *β*-apo-13-carotenal (a C_{18} ketone) (Figure 1.18) (Schwartz *et al*. 2004). The *β*-apo-13-carotenal was then proposed to be converted to strigolactone via an unknown process in the remainder of the pathway. Further *in vitro* experiments by Alder *et al*. with *A. thaliana* and *O. sativa* CCD8 corroborated the findings of Schwartz (Alder *et al*. 2008).

Genetic knockout experiments by Lin *et al*. in *O. sativa* led to the identification of another enzyme on the strigolactone biosynthesis pathway (Lin *et al*. 2009). The enzyme, Dwarf27 (D27, due to the fact that gene knockout mutants display a dwarf phenotype) was shown to be an iron containing enzyme of unknown function which acted upstream of *max*1, a cytochrome P450. The homologue of D27 has subsequently been identified in *Arabidopsis* (Waters *et al*. 2012). D27 does not share sequence similarity with the CCD enzyme family, nor to any other protein of known function. As such, the role of the iron in D27 remained unknown, as did the catalytic mechanism of the enzyme. Experiments by Alder *et al*. have identified D27 as an isomerase, which converts the all-*trans* isomer of *β*-carotene to 9-*cis*-*β*-

carotene (Alder *et al*. 2012). The 9-*cis*-*β*-carotene is then cleaved by CCD7 at the 9',10' position to produce 9-*cis*-*β*-apo-10'-carotenal, which is then converted to a compound named carlactone by CCD8 (Figure 1.19). Carlactone is then converted by the rest of the strigolactone biosynthesis pathway to strigolactone, an unknown process which is known to include at least one P450 oxidase (*max*1) (Booker *et al*. 2005, Lazar & Goodman 2006). The CCD8 reaction is most unusual for a CCD as it appears to catalyse two sequential cleavage reactions on the same substrate.

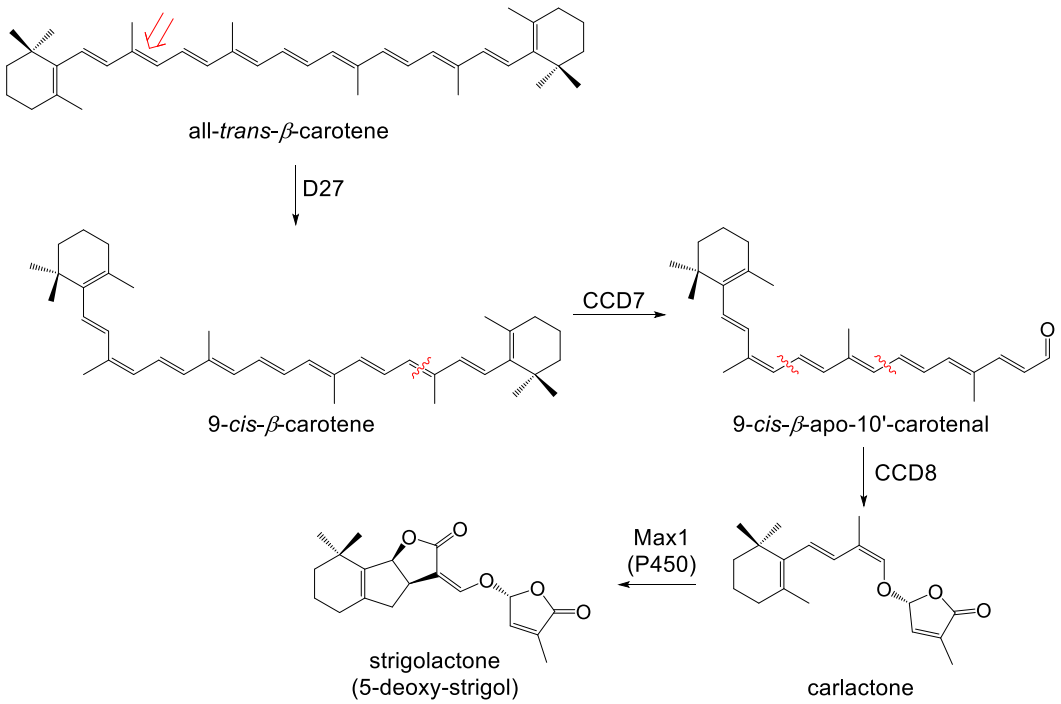


Figure 1.19 – Current proposed biosynthetic route from all-*trans*-*β*-carotene to carlactone, an intermediate on the strigolactone biosynthesis which is converted to strigolactone through an unknown process. The position of isomerisation on all-*trans*-*β*-carotene is indicated with a red arrow. The positions of cleavage in 9-*cis*-*β*-carotene and 9-*cis*-*β*-apo-10'-carotenal are indicated with red lines.

Experiments by Alder and co-workers show that CCD7 cleaves 9-*cis*-*β*-carotene 10 fold faster than it does the all-*trans* isomer. CCD7 has also been shown to cleave the 9-*cis* isomers of zeaxanthin and lutein (Bruno *et al*. 2014). Two groups have also recently reported that the P450 MAX directly catalyses the conversion of carlactone into strigolactone (5-deoxy-strigol) (Zhang *et al*. 2014, Abe *et al*. 2014). Unknown enzymes can then subsequently further oxidise 5-deoxy-strigol to produce other functional strigolactones.

1.5.4 9'-cis-Epoxycarotenoid Dioxygenases (NCEDs) – Abscisic Acid Biosynthesis

Abscisic acid is a plant hormone which is involved in a number of different plant processes such as secondary seed dormancy and drought tolerance (Taylor *et al*. 2005, Iuchi *et al*. 2001, Nambara & Marion-Poll 2005). NCED1 from *Z. mays* was the first CCD enzyme to be biochemically characterised. A vivaporous seed mutant from *Z. mays* was discovered to be deficient in abscisic acid biosynthesis. The mutant was found to contain a lesion in the *vp14* gene, which codes for a protein with sequence similarity to the lignostilbene dioxygenase enzyme family (Schwartz *et al*. 1997). Genetic investigations have shown that, in *A. thaliana* at least, there are five isoforms of the NCED enzyme, termed NCED2, NCED3, NCED5, NCED6 and NCED9. However, the localisation of these isoforms within the plant varies, as does their roles in controlling plant biological processes (Tan *et al*. 2003, Lefebvre *et al*. 2006). *A. thaliana* NCED2 and NCED3 are localised in the roots and reproductive tissues, whereas *A. thaliana* NCED5, NCED6 and NCED9 are localised solely in reproductive tissues. The presence of all NCEDs within the reproductive tissue, and the fact they are expressed during early plant growth is consistent with NCED playing a role in controlling development (Tan *et al*. 2003, Lefebvre *et al*. 2006). In *Arabidopsis*, NCED6 and NCED9 are the most important NCEDs for seed ABA content (Kucera *et al*. 2005). In addition to the control of development, NCEDs control seed germination and plant responses to drought through the action of ABA (Finch-Savage & Leubner-Metzger 2006).

The biosynthesis of abscisic acid begins with 9'-*cis*-neoxanthin, which is produced via a series of enzymatic reactions from *β*-carotene to produce all-*trans*-neoxanthin (Taylor *et al*. 2005, Walter & Strack 2011). An unknown isomerase activity then selectively isomerises the 9',10' double bond in a reaction which appears analogous to that performed by D27 in the strigolactone biosynthesis pathway. NCED then catalyses the 11',12' cleavage of 9'-*cis*-neoxanthin to a C_{25} aldehyde and xanthoxin (Figure 1.20) (Tan *et al*. 1997). The NCED cleavage reaction is the rate limiting step on the abscisic acid biosynthesis pathway. Xanthoxin is then reduced by a short chain alcohol dehydrogenase known as xanthoxin dehydrogenase

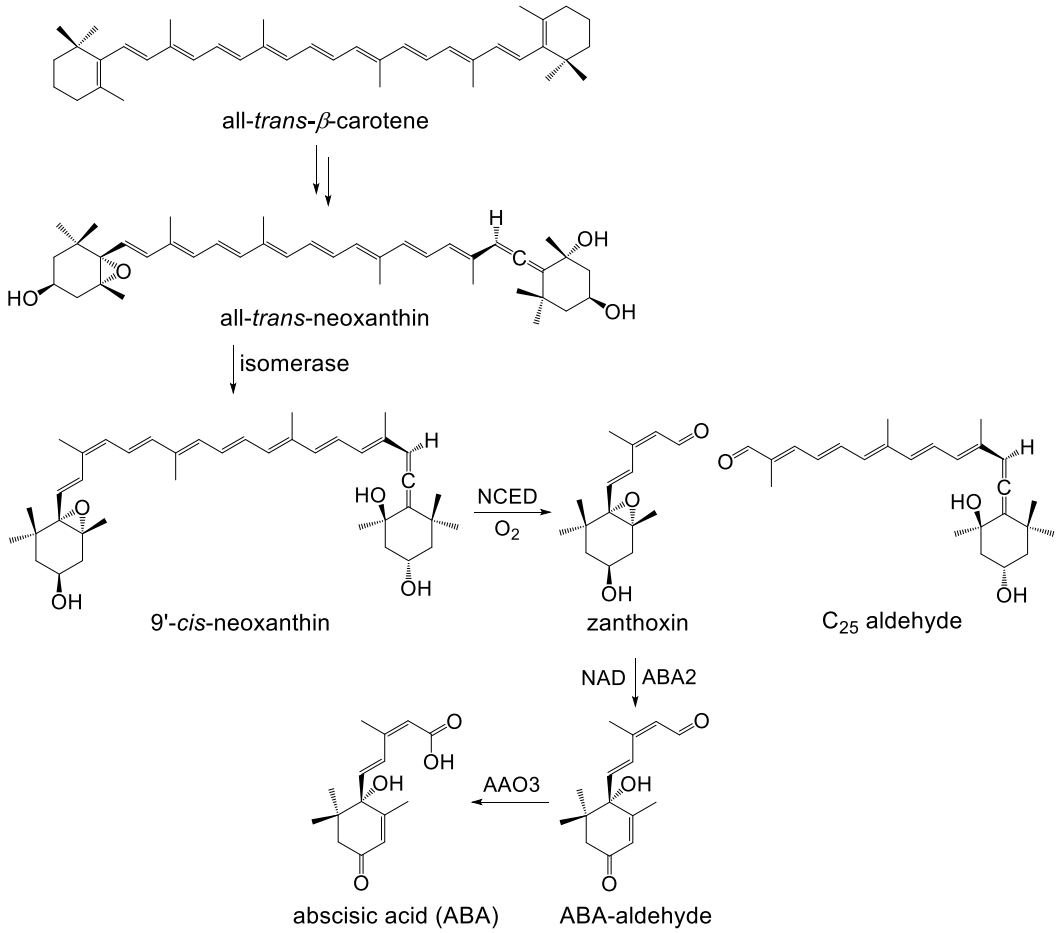


Figure 1.20 – Biosynthetic route from all-*trans*-*β*-carotene to abscisic acid. The rate limiting step, the cleavage of 9'-*cis*-neoxanthin, is catalysed by NCED.

(ABA2) to abscisic aldehyde, which is then finally oxidised to abscisic acid by abscisic acid oxidase (AAO3) (Gonzalez-Guzman *et al*. 2002, Burbidge *et al*. 1999).

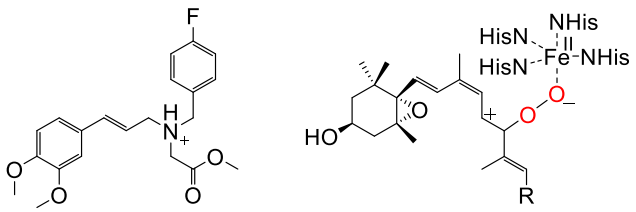
Biochemical investigations with *Z. mays* VP14 have shown that it is selective for the 9-*cis* isomers of the epoxycarotenoids violoxanthin, neoxanthin and luteoxanthin and is targeted to chloroplast membranes in the cell (Tan *et al*. 2001, Schwartz *et al*. 2003). VP14 shows no activity towards the all-*trans* isomers.

1.6 Mechanisms of CCD Cleavage

There is some debate within the literature with regards to the mechanism of cleavage by carotenoid cleavage dioxygenases. Primarily debate surrounds whether CCDs perform a monooxygenase or dioxygenase cleavage of the carotenoid. However, mechanistic studies are limited to a small number of examples performed on *Z. mays* NCED1, mammalian BCO and

A. thaliana CCD1. Early $^{18}O_2$ labelling experiments with *Z. mays* NCED1 suggested a dioxygen mechanism, with almost 100% labelling of abscisic acid in the carboxylic acid position (Zeevaart *et al*. 1989). Subsequent $^{17}O_2$ labelling studies by Leuenberger *et al*. with mammalian BCO gave only 50% incorporation of ^{17}O into the retinal products (Leuenberger *et al*. 2001). Consequentially this was interpreted as a monooxygenase mechanism, contradicting the findings of work on *Z. mays*. Labelling studies using *A. thaliana* CCD1, however, again showed 100% incorporation of ^{18}O into the *β*-ionone and *β*-apo-10'-carotenal products (Schmidt *et al*. 2006). Re-interpretation of the work by Leuenberger suggested that the apparent 50% incorporation on ^{17}O could be due to exchange of isotopically labelled oxygen with unlabelled oxygen in bulk water, as samples were left for 24 hours before analysis. Aldehydes are reactive species and can exchange rapidly with water. Beyond labelling experiments, there is no other experimental data in the literature pertaining to the CCD cleavage mechanism.

It is likely that activation of molecular oxygen in the CCD enzyme reaction occurs via single electron transfer from Fe (II) to oxygen to generate superoxide, as can be deduced by analogy to other non-heme iron dependent enzymes such as extradiol catechol dioxygenases or 2-oxo-glutarate dependent dioxygenases (Bugg 2003, Bugg & Ramaswamy 2008). Attack of superoxide on the carotenoid backbone would create a radical intermediate (Figure 1.21) which would be stabilised via conjugation along the polyene backbone. Subsequent one electron transfer back to the Fe (III) would restore the resting state Fe (II) and create a similarly highly stabilised secondary or tertiary cation, depending upon the position of cleavage. Although no evidence exists for the formation of a cation, it is consistent with inhibition of *Vigna unguiculata* (cowpea) and *A. thaliana* NCED3 by abamine (Figure 1.21 and Section 1.8). At physiological pH, the tertiary amine of abamine is likely to be protonated, resulting in a tertiary amine cation. This cation would mimic the substrate cation and bind to the active site, inhibiting the cleavage reaction.



abamine secondary substrate cation

Figure 1.21 – Structures of the protonated abamine and the secondary substrate cation formed during carotenoid cleavage.

Following formation of the secondary cation, one of three reactions is possible (Figures 1.22 and 1.23). Firstly, the oxygen anion could ring close onto the cation to form a highly strained four membered dioxetane ring (path one). Breakdown of the dioxetane would result in the two aldehyde products (Schmidt *et al.* 2006, Harrison & Bugg 2013). The second possible mechanism involves nucleophilic attack of water on the carbocation (path two) followed by a Criegee rearrangement. This would result in a secondary carbocation which would be quenched by iron (II) hydroxide. A similar mechanism is proposed to occur in the extradiol catechol dioxygenases and the resulting hemi-acetal would break down to form the two aldehyde products (Bugg 2003). A third possible mechanism, the monooxygenase mechanism (Figure 1.23), also exists, whereby an epoxide forms following formation of the carbocation (Leuenberger *et al*. 2001). Formation of the epoxide would result in a high energy iron (IV) intermediate and attack on the epoxide by water would result in the formation of an α,β-diol, which breaks down to form the two aldehyde products.

Without detailed mechanistic data it is very difficult to distinguish between the possible mechanisms. Labelling experiments with $^{18}O_2$ strongly suggest a dioxygen mechanism, incorporating both atoms of oxygen from O_2. However, whether that mechanism proceeds through a dioxetane intermediate or Criegee intermediate is unknown. It was noted in the crystal structure of VP14 that there is a lack of acidic and basic residues within the active site, which would lend support for a dioxetane mechanism (Messing *et al*. 2010). However, in the case of 9,10 cleavage reactions, formation of the dioxetane would require attack on a tertiary cation, which could present a steric challenge for the enzyme. In addition to this, dioxetane formation in catalysis is very rare. Very few enzymes are known to proceed

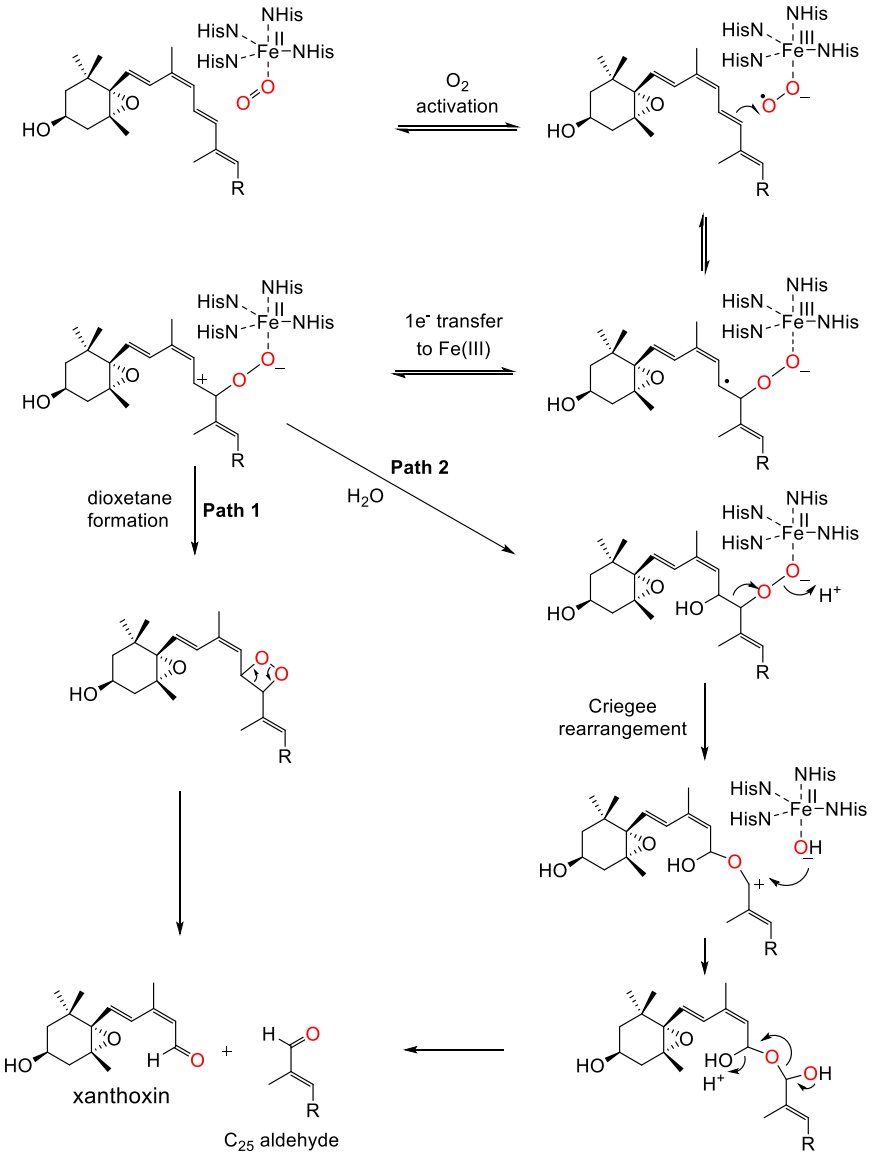


Figure 1.22 – Proposed dioxygen cleavage mechanisms for CCDs based upon the NCED catalysed cleavage of 9'-*cis*-neoxanthin. Path one involves the formation of a dioxetane intermediate which breaks down to form two aldehyde products. Path two, however, utilises a Criegee rearrangement (which can also proceed in the reverse direction). Figure adapted from Harrison & Bugg 2013.

through a dioxetane mechanism as formation of a dioxetane ring is an unfavourable endothermic process. Vitamin K-dependent carboxylase and firefly luciferase are two examples of enzymes whose mechanisms are known to proceed through dioxetane intermediates (Suttie 1985, Dowd *et al*. 1994, McCapra 1976, Franks & Brick 1996).

The Criegee rearrangement mechanism, however, requires the addition of water, so $^{18}O_2$ experiments would result in either one or two atoms of ^{18}O being incorporated depending upon the direction of the final C-O cleavage step. Only the dioxetane mechanism would

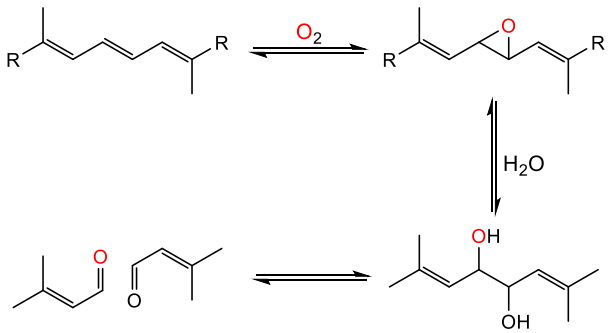


Figure 1.23 – Possible monooxygenase cleavage mechanism of carotenoids by CCDs, proposed by Leuenberger *et al*. (Leuenberger *et al*. 2001). Figure from Harrison & Bugg 2013.

guarantee incorporation of both atoms from molecular oxygen. Another issue is the source of water to attack the substrate cation, which could be provided via the semi conserved second shell glutamic acid residues participating in catalysis, possibly by deprotonating water. This could arise through the iron cofactor acting as a Lewis acid. The sixth ligand on the iron (II), most likely water, would have its pK_a lowered to between 7-10. A proton this acidic could then be abstracted by one of the second shell glutamate residues, allowing participation in catalysis. A metal cofactor acting as a Lewis acid is not unprecedented. In the carboxypeptidase active site the zinc cofactor decreases the pK_a of bound water allowing deprotonation by a glutamate ligand (Christianson & Lipscomb 1989). If this were the case it would lend support for the Criegee rearrangement.

As discussed, there is a high degree of sequence similarity between the carotenoid cleavage dioxygenases and the lignostilbene dioxygenases. Labelling studies using $^{18}O_2$ on two lignostilbene dioxygenases identified from *Novosphingomonas aromaticivorans* have shown only 50% incorporation of the ^{18}O label into the products, suggesting a monooxygenase mechanism for these lignostilbene dioxygenases (Marasco & Schmidt-Dannert 2008). In addition, the LSDs assayed showed no activity against carotenoids or apocarotenoids. The inability of LSD to cleave carotenoids is unsurprising, given how CCDs tend to cleave substrates of similar structures (Wirtz *et al*. 2001). However, the observation of a monooxygenase mechanism is interesting. It is entirely possible that the reaction is proceeding via a dioxygenase mechanism, but with a Criegee rearrangement, which eliminates the other ^{18}O label. Such a mechanism would explain why only 50% labelling is observed.

Computational studies have been performed examining the thermodynamics of possible CCD cleavages (Borowski *et al*. 2008). Such studies, which examine the lowest energy routes through the reaction, suggest that energetically the dioxetane intermediate is more likely than the epoxide intermediate. However, the study did not examine the possible Criegee rearrangement, so such a mechanism cannot be discounted.

Little insight into carotenoid cleavage is provided by model chemistry. To date, only one model complex, involving a ruthenium porphyrin, has been demonstrated (French *et al*. 2000). The ruthenium complex has been shown to cleave *β*-carotene to two equivalents of retinal in the presence of *t*-butyl hydroperoxide.

1.7 Carotenoid Isomerisation

In addition to carotenoid cleavage activity, another role of CCDs has emerged; the isomerisation of carotenoid substrates. The enzyme RPE65, which shares sequence similarity to the CCD enzyme family, although not strictly a CCD, isomerises all-*trans*-retinyl esters to 11-*cis* retinol (Oberhauser *et al*. 2008). Debate surrounds the proposed mechanism of isomerisation by RPE65. It was initially proposed by Deigner *et al*. (before the discovery of RPE65) that isomerisation occurred through an enzyme active site nucleophile specifically attacking the 11,12 double bond of all-*trans*-retinyl ester, eliminating a carboxylate (Deigner *et al*. 1989). This removes the double bond character at the 11,12 double bond allowing isomerisation (Figure 1.24). Attack on the terminal double bond by water eliminates the active site base and forms the 11-*cis*-retinal. Several issues with this mechanism were raised by McBee *et al*. including the breaking a highly stable conjugated double bond system through nucleophilic attack and addition of water to conjugated double bond system (McBee *et al*. 2000).

A second mechanism involving a carbocation mediated mechanism was proposed by McBee *et al*. (McBee *et al*. 2000). Here retinol is protonated by an active site acid, resulting in elimination of water and the formation of a primary cation, which is delocalised along the

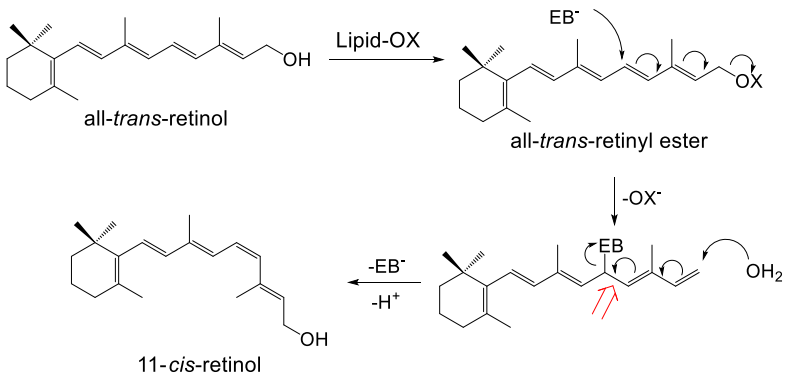


Figure 1.24 – Proposed isomerisation mechanism of all-*trans*-retinol to 11-*cis*-retinal involving attack of an active site base specifically on the 11,12 double bond of all-*trans*-retnyl-ester, resulting in the loss of polyene double bond character, allowing isomerisation. Attack of water on the terminal alkene reforms the polyene system and eliminates the active site base.

conjugated system. This results in a loss of double bond character, allowing isomerisation. Addition of water to the cation re-establishes the conjugated system and forms the 11-*cis*-retinol (Figure 1.25). Further evidence in support of the cation mediated mechanism was provided by Redmond *et al*. following the discovery of RPE65 (Redmond *et al*. 2010). It was observed *in vitro* that RPE65 was a leaky enzyme, producing the 13-*cis* isomer of retinal in addition to the 11-*cis* isomer, which is consistent with a carbocation mediated mechanism. Mutations which removed aromaticity in the active site all resulted in a decrease in activity, consistent with a loss of cation stabilisation.

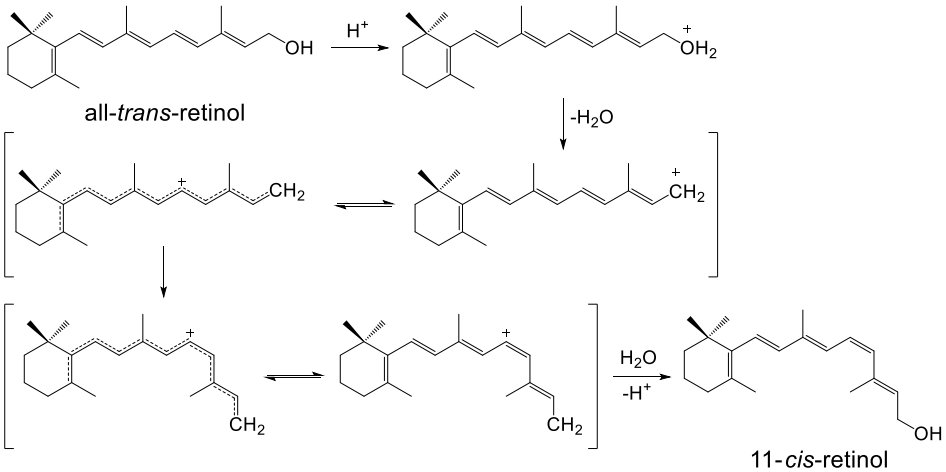


Figure 1.25 – Proposed isomerisation mechanism of all-*trans*-retinol to 11-*cis*-retinal involving the formation of a substrate cation, resulting in a loss of double bond character, allowing free rotation of the delocalised polyene backbone.

^{18}O labelling studies by Kiser *et al*. with H_2O^{18} revealed that the oxygen on the retinol originated from bulk water (Kiser *et al*. 2009). Consequentially it was proposed that isomerisation of the all-*trans*-retinyl ester by RPE65 occurs through an *O*-alkyl cleavage reaction (Figure 1.26) (Kiser *et al*. 2012). The ferrous iron acts as a Lewis acid, coordinating the ester moiety and delocalising electron density towards the carbonyl. This results in cleavage of C-O of the ester, forming a primary carbocation which is delocalised along the polyene backbone. Delocalisation of the cation results in a loss of double bond character, allowing free rotation of the of the C-C bonds. Attack of water on the terminus of the carotenoid cation results in formation of 11-*cis*-retinal. Although unusual, iron is known to act as a Lewis acid, a feature which is likely enhanced by coordination to four uncharged histidine ligands (Sanvoisin *et al*. 1995). Aromatic residues within the active site have been shown to mediate the isomerisation reaction to ensure the formation of the 11-*cis* product and modulate the progression of the reaction (Chander *et al*. 2012).

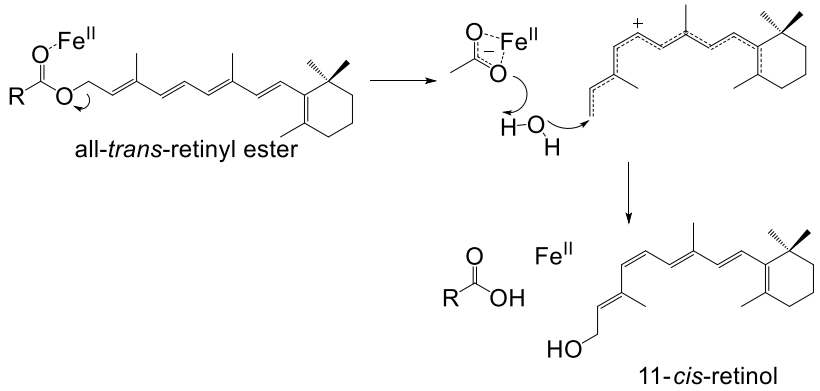


Figure 1.26 – Current accepted mechanism for the isomerisation of all-*trans*-retinyl ester to 11-*cis*-retinol in which ferrous iron acts a Lewis acid, causing ester hydrolysis. Free rotation can occur in the subsequent cation before attack of water forms the 11-*cis*-retinol product.

Isomerisation has also been observed in the CCD NinaB from *G. mellonella*, which not only cleaves carotenoids but also forms isomerised cleavage products (Oberhauser *et al*. 2008). In this case there is no ester moiety with which the carotenoid substrate can bind to the iron. As a result, the question arises as to how is NinaB able to catalyse an isomerisation reaction? Examination of the carotenoid cleavage mechanism provides some possible answers (Figure 1.27) (Harrison & Bugg 2013). Following the formation of the substrate radical or

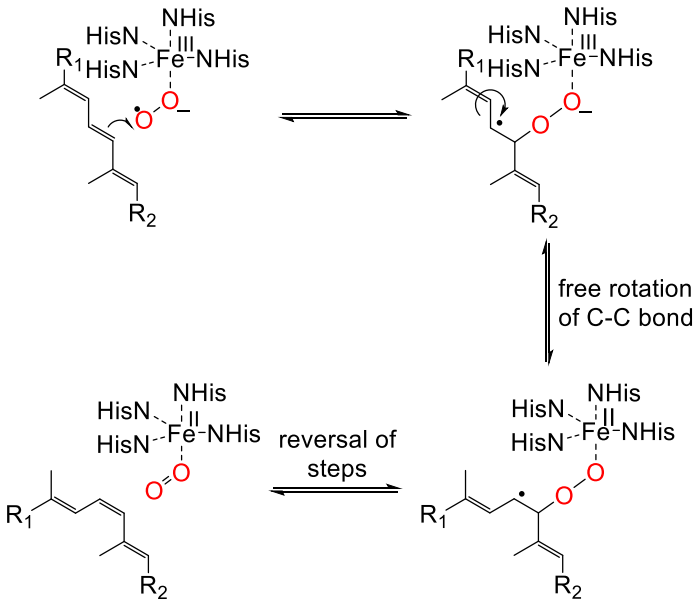


Figure 1.27 – Possible mechanism for the isomerisation of carotenoids. Following formation of the secondary radical, double bond character is lost, allowing free rotation and isomerisation. Reversal of the steps would then result in the isomerised carotenoid product.

cation the double bond character of the carotenoid is lost. Just as in the case of the RPE65 isomerisation, at this point free rotation of the C-C bonds on the backbone is allowed. The reaction can then either proceed to completion, resulting in the formation of the two isomerised oxidised products, or reverse, since formation of the substrate radical is reversible thus forming the isomerised carotenoid. Radical mediated isomerisation is precedented in non-heme iron (II) dioxygenase chemistry. The *cis-trans* isomerisation of cycloproyl analogues was observed by Spence *et al*. with the extradiol cleavage dioxygenases MhpB from *Escherichia coli* and *Alcaligenes eutrophus* (Spence *et al*. 1996). In the case of NinaB it is likely that isomerisation occurs during the cleavage reaction.

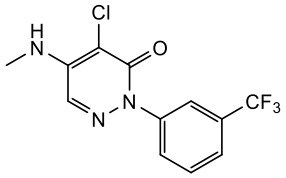
It was suggested by Kloer *et al*. that the *Synechocystis* sp. ACO was also able to perform isomerisation reactions (Kloer *et al*. 2005). Electron density observed within the active site was assumed to belong to an isomerised carotenoid substrate. However, this electron density has since been shown not to be due to a bound carotenoid substrate and ACO has been shown not to catalyse isomerisation reactions (Sui *et al*. 2014).

Two other enzymes not belonging to the carotenoid cleavage dioxygenases enzyme family have also been shown to catalyse carotenoid cleavage. The first of these to be identified was prolycopene isomerase (CRTISO), first identified in *S. lycopersicum* and subsequently in *A. thaliana* (Isaacson *et al*. 2002, Park *et al*. 2002). Prolycopene isomerase is responsible for the isomerisation of 7,9/7',9'-*cis*-phytoene to all-*trans*-phytoene, the final step in the biosynthesis of all-*trans*-phytoene in plants. Prolycopene isomerase has been shown to contain a flavin adenine dinucleotide (FAD(H)) or nicotinamide adenine dinucleotide (NAD(H)) cofactor binding domain (Isaacson *et al*. 2004). Prolycopene isomerase from *S. lycopersicum* activity has been found to be dependent upon reducing conditions and upon the presence of cytoplasmic extracts. CTRISO is inactive in the purified form without the addition of cytoplasmic components from lysed *E. coli* (Isaacson *et al*. 2004) It has been suggested that CRTISO receives an electron from complex I (NADH ubiquinone oxidoreductase) in the electron transport chain and uses this electron to break the conjugation of 7,9/7',9'-*cis*-phytoene, thus allowing isomerisation. The electron is then passed onto complex II (succinate ubiquinone oxidoreductase). This proposed mechanism is consistent with the observation that CRTISO is dependent upon other cytosolic materials, however, it does not explain the function of the FAD(H)/NAD(H) binding domain or suggest how electron transfer is mediated by CRTISO.

The second enzyme, Dwarf27 first identified in *O. sativa* and subsequently in *A. thaliana*, is required for the isomerisation of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene in the biosynthesis of strigolactone (see Section 1.5.3) (Lin *et al*. 2009, Alder *et al*. 2012, Waters *et al*. 2012). Interestingly D27 shows no sequence similarity to prolycopene isomerase, nor does it contain an FAD(H) binding domain like prolycopene isomerase. D27 has, however, been shown to contain 1.7 moles of iron for every one mole of protein. Currently the mechanisms of isomerisation by both prolycopene isomerase and Dwarf27 remain unknown.

1.8 Inhibition of Carotenoid Cleavage Dioxygenases

Within the literature there are several different examples of small molecule compounds which inhibit carotenoid biosynthesis and subsequently apocarotenoid production (Bramley 1993). One of the most well-known compounds is norflurazon (Figure 1.28), which inhibits the enzyme phytoene desaturase (Bartels & Watson 1978). However, these compounds all act high up on the carotenoid biosynthesis pathway. As a result, these compounds are non-specific, inhibiting almost all carotenoid and apocarotenoid mediated processes. However, over the last decade, several different CCD specific inhibitors have been identified. The first CCD specific inhibitor identified was nordihydroguaiaretic acid (NDGA), originally developed as a lipoxygenase inhibitor which is known to inhibit ABA biosynthesis (Creelman *et al*. 1992). NDGA was shown to inhibit *V. unguiculata* NCED1 *in vitro* by 80% at 100 µM (Han *et al.* 2004a, Han *et al*. 2004b). Given the structure of NDGA (Figure 1.29), it is likely it inhibits by acting as a substrate analogue. Binding to the active site and blocking carotenoid access.



norflurazon

Figure 1.28 – Chemical structure of norflurazon.

A second NCED inhibitor was subsequently developed, abamine (Figure 1.29) (Han *et al.* 2004a). Abamine is a tertiary amine compound with a K_i of 38.8 µM *in vitro* against *V. unguiculata* NCED1. At physiological pH, it is likely that the tertiary amine of abamine will be protonated. This tertiary amine cation could then inhibit the CCD enzyme by mimicking the cation which is believed to form during the cleavage of carotenoids by CCDs (see Section 1.6). Interestingly abamine shows some selectivity towards NCEDs. Abamine has been shown to inhibit *S. lycopersicum* CCD1 *in vitro* by 35% at 100 µM (Sergeant *et al*. 2009). A second generation abamine inhibitor, known as abamineSG, has been developed (Figure 1.29) (Kitahata *et al*. 2006). AbamineSG is very similar to abamine, however, it contains a longer

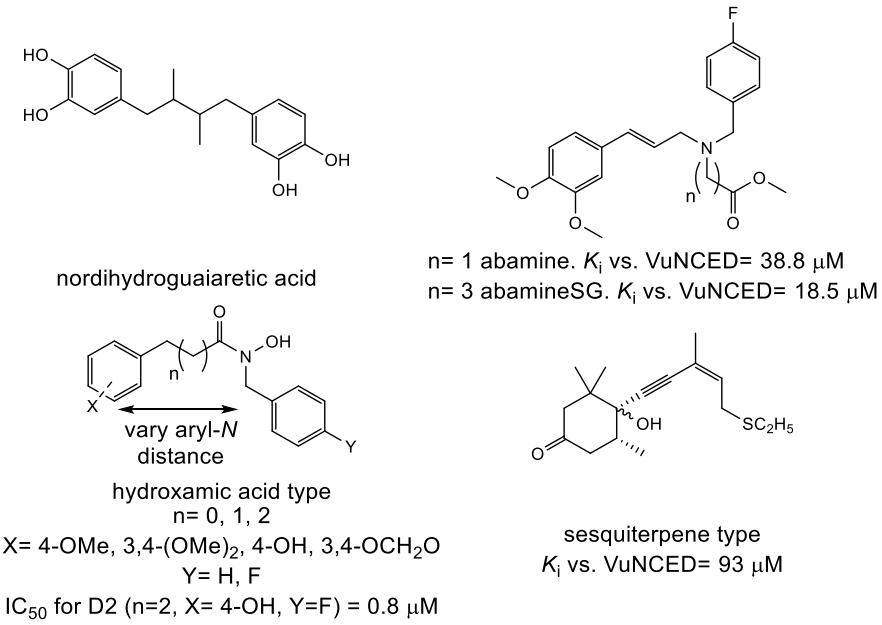


Figure 1.29 – Structures of the four classes of small molecule CCD inhibitors.

ester moiety, most likely increasing the affinity with which the molecule can bind to the hydrophobic active site with. *In vitro* against *V. unguiculata* NCED1 a K_i of 18.5 μM was reported. *In planta* both abamine and abamineSG inhibit ABA biosynthesis.

A family of sesquiterpene like compounds have also been shown to inhibit CCD cleavages and lead to reductions of ABA levels *in planta* (Figure 1.29) (Boyd *et al*. 2009). Boyd *et al*. have reported almost 100% inhibition of *A. thaliana* NCED3 *in vitro* by some sesquiterpene like inhibitors at 100 μM. However, the sesquiterpene like compounds are less effective than abamine and abamineSG, with the most effective compound having a modest K_i of 93 μM. Being sesquiterpene like compounds, these inhibitors are proposed to inhibit NCED via a mechanism similar to that of NDGA.

Despite several compounds being shown to be effective inhibitors of NCED, only one class of compounds have been shown to inhibit other CCD enzymes. Hydroxamic acids are known iron chelators which have been shown to be effective at inhibiting *S. lycopersicum* CCD1a. *In vitro* at 100 μM, 100% inhibition is seen by some members of the hydroxamic family, and IC_{50} values lower than <1 μM are reported (Sergeant *el al*. 2009). The hydroxamic acids are believed to inhibit CCDs by binding to the active site iron, blocking carotenoid

access to the active site. It is proposed that by varying the aryl-*N* length (Figure 1.29), the specificity of the hydroxamic acids can be altered (Figure 1.29 and see Chapters 2 and 3 for a complete list of hydroxamic acid inhibitors). However, against *S. lycopersicum* NCED1 *in vitro* at most 33% inhibition at 100 µM was observed. This is somewhat lower than the inhibition seen for abamine, abamineSG and the sesquiterpene like compounds, and also lower than the inhibition seen against CCD1. Differences in active site architecture could well explain the differences in the levels of inhibition observed. *In vivo* against *A. thaliana* CCD7 weak inhibition (less than 50% inhibition at 100 µM) has been seen by the hydroxamic acids. However, low inhibition here may be due to poor uptake of the inhibitor by *E. coli*. Additionally, against *M. musculus* BCOI and BCOII *in vivo* at 100 µM less than 30% inhibition was recorded. Given that BCOI is a 15,15' cleavage enzyme, a lack of inhibition is expected. However, BCOII has 9,10/9',10' *β*-carotene cleavage specificity like CCD1. As a result, inhibition levels similar to CCD1 would be expected for BCOII. This could suggest a different mechanism of inhibition. However, poor uptake into *E. coli* is the likely cause of the poor inhibition observed.

In planta, phenotypic effects on shoot branching and seed germination have been seen in the presence of several members of the hydroxamic acid inhibitor family (Figure 1.30) (Sergeant *et al*. 2009, Sergeant *et al*. unpublished). For example, the compound D4 (Table 1.1) when applied to seeds can partially overcome the osmotic inhibition of germination. Effects on seed germination are believed to be mediated through inhibition of NCED. NCED is the rate limiting step in the biosynthesis of ABA, which induces secondary seed dormancy, preventing seed germination (Finch-Savage & Leubner-Metzger 2006). Effects on shoot branching have also been observed (Sergeant *et al*. 2009). When compound D6 (Table 1.1) is applied to wild type *A. thaliana*, there is a two fold increase in the number of shoot branches produced (Figure 1.30). Mediation of the shoot branching phenotype is thought to occur through inhibition of CCD7 and/or CCD8, which controls the biosynthesis of strigolactone, a phytohormone which inhibits shoot branching in plants (Gomez-Roldan *et al*. 2008, Umehara

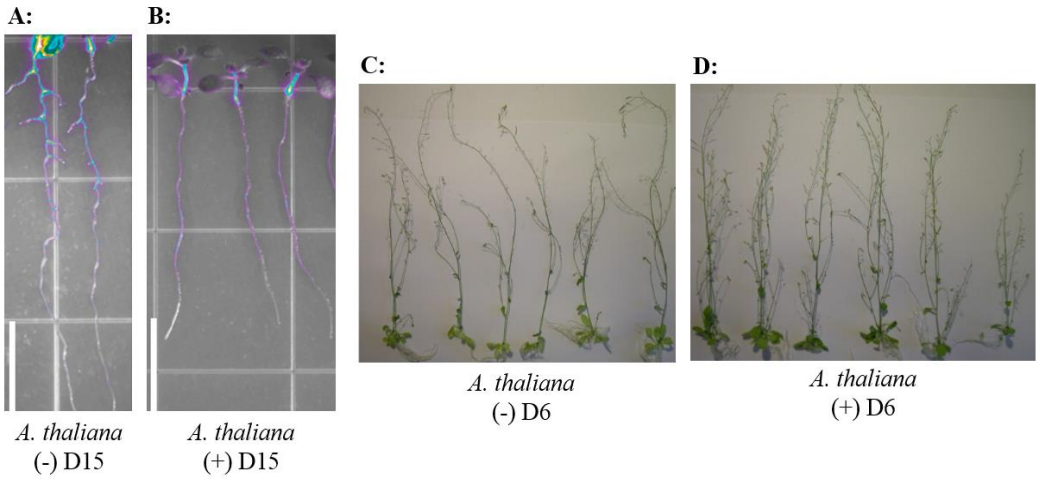


Figure 1.30 – *In planta* phenotypes observed in *A. thaliana* on addition of hydroxamic acid inhibitors. Figures A and B show lateral root branches in wild type (Columbia-0) *A. thaliana* without (**A**) and with (**B**) the application of 125 µM D15. Absence of lateral roots in the presence of D15 suggest inhibition of a carotenoid derived molecule. Figures C and D show shoot branching in wild type (Columbia-0) *A. thaliana* without (**C**) and with (**D**) 100 µM D6. Increased shoot branching in the presence of compound D6 suggests inhibition of strigolactone biosynthesis. Images from Van Norman *et al*. 2014 and Sergeant *et al*. 2009.

et al. 2008). Additionally, with the hydroxamic acids F1 and F2 (Table 1.1), phytotoxic effects on *A. thaliana* have been observed which can be observed through bleaching of the plants (Figure 1.31) (Sergeant *et al*. 2013). This phytotoxicity effect has been shown to occur partially through inhibition of metalloenzyme *p*-hydroxy-phenylpyruvate dioxygenase, a known herbicide target (Sergeant *et al*. 2013). However, studies indicate that a second, unknown enzyme, is also involved. One hydroxamic acid, D15 (Table 1.1), has also been shown to effect lateral root branching in *A. thaliana* (Figure 1.30) (Van Norman *et al*. 2014). The biochemical target of this inhibitor is unknown and is believed to constitute a new role for carotenoids *in planta*.

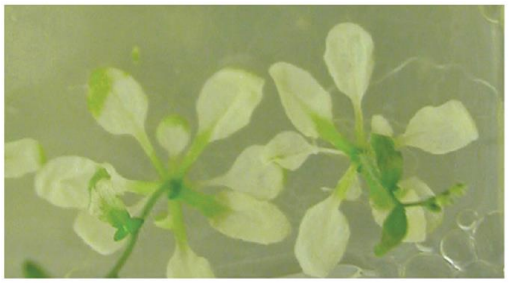


Figure 1.31 – Bleaching observed on wild type (Columbia-0) *A. thaliana* in the presence of 100 µM compound F1. Image from Sergeant *et al*. 2013.

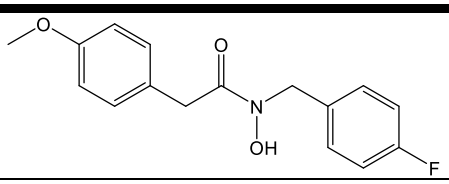
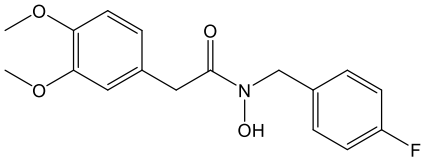
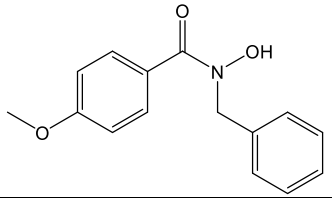
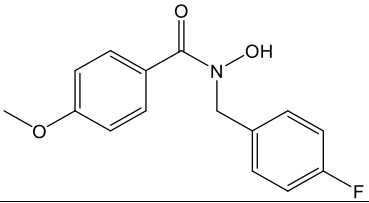
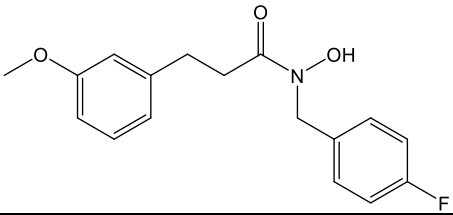
Phenotype	Species Observed in	Mediated by?	Inhibitor	Structure
Seed Germination	*Arabidopsis thaliana, Z. mays*	NCED	D4	
Shoot Branching	*Arabidopsis thaliana*	D27 or CCD7 or CCD8	D6	
Bleaching	*Arabidopsis thaliana*	HPPD and unknown target	F1	
			F2	
Lateral root Branching	*Arabidopsis thaliana*	Unknown	D15	

Table 1.1 – Summary of phenotypic effects observed *in planta* on addition of hydroxamic inhibitors to *A. thaliana* and *Z. mays* and the enzymes believed to be involved in mediation of the phenotype.

Inhibitors of lignostilbene dioxygenases have been reported. Han *et al*. have shown that (*Z*)-1-(4-hydroxyphenyl)-1-fluoro-2-phenylethene (Figure 1.32) is a competitive inhibitor *in vitro* of lignostilbene-*α*,*β*-dioxygenase (LSD) from *P. paucimobilis*, with a K_i of 0.98 μM (Han *et al*. 2002). Effective inhibition of LSD was found to result from polarisation of the double bond, with the fluorine in the *Z* configuration, which limits delocalisation of electrons. Additionally, a 4-hydroxy group was also required for inhibition and is proposed to be essential for binding of the substrate with the active site of the enzyme. However, when tested *in vitro* against *A. thaliana* NCED3, (*Z*)-1-(4-hydroxyphenyl)-1-fluoro-2-phenylethene showed no inhibition, which is proposed to be due to a lack of binding of the inhibitor to the NCED active site. Han *et al.* have also reported the inhibition of LSD by *N*-benzylideneaniline

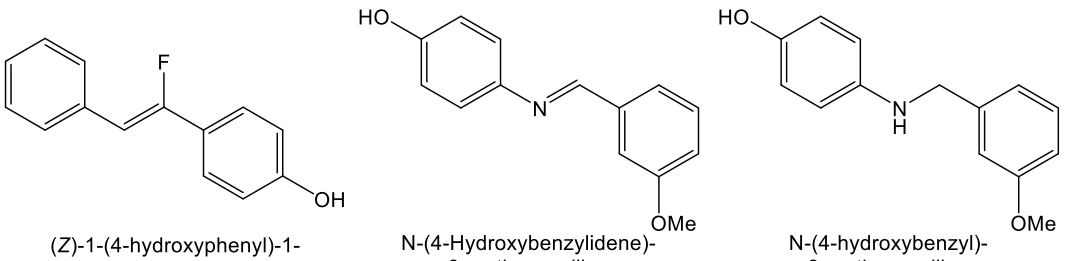


Figure 1.32 – Structures of the lignostilbene dioxygenases inhibitors (*Z*)-1-(4-hydroxyphenyl)-1-fluoro-2-phenylethene, *N*-(4-Hydroxybenzylidene)-3-methoxyaniline and *N*-(4-hydroxybenzyl)-3-methoxyaniline

and *N*-benzylaniline derived compounds (Han *et al*. 2004b). *In vitro* against *P. paucimobilis* LSD the compounds *N*-(4-Hydroxybenzylidene)-3-methoxyaniline and *N*-(4-hydroxybenzyl)-3-methoxyaniline (Figure 1.33) were shown to have IC_{50} values of 0.3 μM and 10 μM respectively. However, neither compound was effective *in vitro* against *A. thaliana* NCED3.

1.9 Aims

By 2050 it is projected that the world population will have reached 9.5 billion people, straining world food production resources (United Nations World Population Prospects, 2012 Revision). It is therefore of paramount importance to maximise land use efficiency, ensuring high yielding crops are produced. Many of the factors which can lead to crop failure or low yields are due to abiotic stress (non-living factors), such as temperature and water levels (Mendelsohn 2007). Abiotic stress management seeks to control these factors to ensure high crop yields. For example, secondary seed dormancy in plants is controlled by the hormone abscisic acid, the levels of which in seeds can vary according to several factors including osmotic potential and temperature (Finch-Savage & Leubner-Metzger 2006). It is proposed that abscisic acid levels in seeds can be controlled through a chemical genetics approach which lowers abscisic acid levels ensuring crops could be planted with a high certainty of germination, reducing crop failure and increasing yields.

CCD enzymes are vital for the biosynthesis of the phytohormones abscisic acid and strigolactone which control many important plant process. As a result, much work has been conducted in order to understand the biochemical roles that CCDs play in different organisms.

One such method, in addition to knock-out mutants, is to use inhibit enzymes with small molecules, such as the inhibitors described in Section 1.8. In order to understand the pathways which control the biosynthesis of these important phytohormones, a detailed understanding of the targets of the small molecule interventions and the mechanisms of the enzymes is required. In particular, the mechanisms of carotenoid cleavage dioxygenases, which catalyse the rate limiting steps in strigolactone and abscisic acid biosynthesis.

The aim of this project, the results of which are described in the coming chapters, is to understand the biochemical basis for the phenotypic effects observed. Previously published studies have only studied the tomato CCD1a and NCED1 enzymes *in vitro* (Sergeant *et al.* 2009). Expanding the *in vitro* study to explore the enzymes of the strigolactone biosynthesis pathway and to study seed germination phenotypes observed in more detail will be of benefit for the development of potential agrochemical compounds. *In planta* data detailed effects on seed germination and shoot branching in the presence of hydroxamic acids does not completely correlate with *in vitro* data. Understanding the biochemical basis of these phenotypes, the mechanisms of the enzymes involved along with detailed structure-activity relationships, is required in order to develop new effective inhibitors. A selective inhibitor for each of the CCD enzymes would aid in understanding the molecular basis for the seed germination and shoot branching activities.

The specific aims of the project are as follows:

- CCD1:
    - Obtain structure activity relationship data of hydroxamic acid inhibitors versus *S. lycopersicum* CCD1a.
    - Develop a method for the purification of *S. lycopersicum* CCD1a.
- NCED:
    - Obtain structure activity relationship data of hydroxamic acid inhibitors versus *S. lycopersicum* NCED1.

- Express and assay *Z. mays* NCED1, 2, 3A, 3B and 9 versus compounds D2 and D4.
- Clone and express *A. thaliana* ABA2 enzyme and develop a high throughput assay for the screening of potential inhibitors against NCED.

- D27 / CCD7:
  - Obtain structure activity relationship data of hydroxamic acid inhibitors versus *A. thaliana* CCD7.
  - Clone and express the D27 isomerase from *O. sativa*, perform structure activity relationship analysis versus the hydroxamic acid inhibitors and perform initial mechanistic investigations.
- CCD8:
  - Obtain structure activity relationship data of hydroxamic acid inhibitors versus *A. thaliana* CCD8.
  - Investigate the mechanism of *A. thaliana* CCD8.

Chapter Two: Expression, Purification and Inhibition of *Solanum lycopersicum* Carotenoid Cleavage Dioxygenase 1a

2.1 Introduction

First identified in *Arabidopsis thaliana* (thale cress), carotenoid cleavage dioxygenase 1 (CCD1) is a cytosolic CCD enzyme responsible for the production of the volatiles *β*–ionone and geranylacetone (Figure 2.1) (Schwartz *et al*. 2001). Many orthologs of *A. thaliana* CCD1 have been identified in many different plant species including *Solanum lycopersicum* (tomato), *Rosa damascena* (rose) and *Cucumis melo* (melon) (Simkin *et al*. 2004, Huang *et al*. 2009, Ibdah *et al*. 2006). CCD1 has been shown to accept a wide variety of different carotenoid, epoxycarotenoid and apocarotenoid substrates, such as *β*–carotene, lycopene and *β*-apo-8'-carotenal (Auldridge *et al*. 2006, Schwartz *et al*. 2001). Despite this high substrate promiscuity, CCD1 shows a high selectivity for the 9,10/9',10' double bond of these substrates both *in vitro* and *in planta*. *In vitro* experiments have also shown the formation of products resulting from a 5,6/5,6' cleavage of lycopene. However, these products have not been observed *in vivo* or *in planta* (Vogel *et al*. 2008).

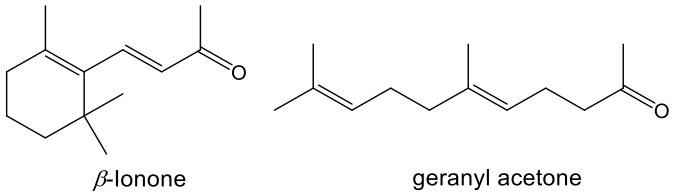


Figure 2.1 – Structures of the aroma volatiles *β*-ionone and geranyl acetone.

Due to its substrate promiscuity, CCD1 can be assayed via a continuous assay with the substrate all-*trans*-*β*-apo-8'-carotenal (Figure 2.2) (Sergeant *et al*. 2009). Although the cytosolic nature of CCD1 makes it easier to study than its plastid located cousins, many authors have found purification of CCD1 difficult due to instability and the formation of

inclusion bodies (Schmidt *et al*. 2006). In their 2009 paper, Sergeant *et al.* report the assay of several hydroxamic acid based inhibitors against *S. lycopersicum* CCD1 (Sergeant *et al.* 2009). However, a complete structure activity relationship study of the hydroxamic acids against LeCCD1a has not been performed.

This chapter describes work on the development of novel purification techniques for LeCCD1a. In addition, the structure activity relationship of the hydroxamic acids towards LeCCD1a is also detailed.

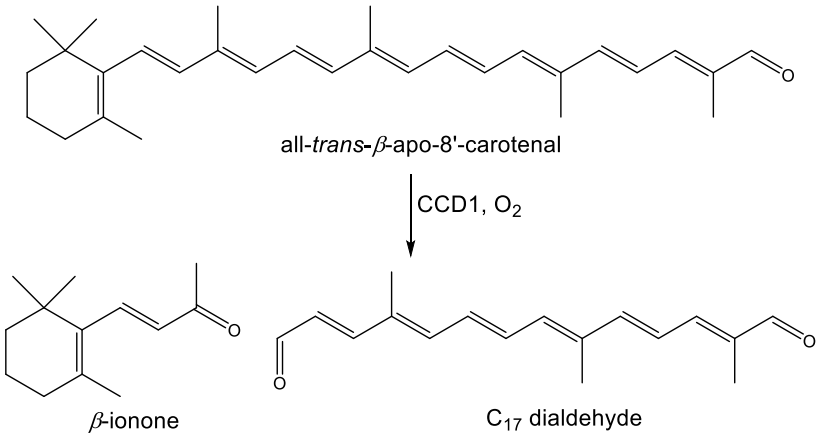


Figure 2.2 – Cleavage reaction performed by CCD1 on the all-*trans*-*β*-apo-8'-carotenal substrate.

2.2 Purification and Assays of *S. lycopersicum* CCD1a

Little biochemical information on the CCDs has been reported in the literature, with the bulk of knowledge on the CCD enzyme family coming from *in vivo* and *in planta* studies. The successful purification of CCD enzymes by affinity chromatography such as glutathione S-transferase (GST) tagged purification and N-His_6 tagged purification has been reported (Alder *et al*. 2008, Messing *et al*. 2010). Some members of the CCD1 enzyme family, such as *A. thaliana* CCD1, can form inclusion bodies during purification, thus leading to a loss of activity of the purified enzyme (Schmidt *et al*. 2006). This may be due to the presence of hydrophobic patches on the surfaces of CCD enzymes which are required for association with the lipid membranes, as demonstrated with the crystal structures of *Synechocystis sp.* apocarotenoid cleavage oxygenase (ACO) and *Zea mays* VP14 (Kloer *et al*. 2005, Messing *et al*. 2010). In addition, by analogy to other dioxygenases such as the catechol dioxygenases,

the catalytic cycle of the CCD is likely to involve highly reactive species such as oxygen radicals (Bugg 2003, Bugg & Ramaswamy 2008). These intermediates can act non-specifically to attack residues of the enzyme active site, resulting in denaturation of the enzyme and causing a loss of enzyme activity. Air oxidation of the catalytic iron (II) over time to iron (III) could also result in a loss of enzyme activity if the cleavage mechanism begins with iron (II), as is predicted (Harrison & Bugg 2013).

Two LeCCD1a plasmids were provided by Dr Andrew Thompson (Cranfield University), pGEX-4T-1-LeCCD1a (GST-LeCCD1a) and pET-14b-LeCCD1a (N-His_6-LeCCD1) (accession number AY576001). Initial experiments into the purification of LeCCD1a were attempted via glutathione-S-transferase tag purification. GST tag purification was used in preference to N-His_6 tag purification due to the presence of the iron cofactor within the enzyme active site. However, the GST tag is rather bulky, adding a molecular weight of approximately 30.0 kDa to the overall mass of the protein being purified, which risks interfering sterically with the catalytic activity of the enzyme.

The pGEX-LeCCD1a plasmid was used to transform pRosetta BL21 *E. coli* and the LeCCD1a enzyme was overproduced by addition of 1 mM IPTG followed by growth for 16 hours at 16º C (Sergeant *et al.* 2009). In order to help overcome non-specific binding of cytosolic components, GST resin was incubated for at least one hour with cell lysate at 4º C. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of eluted fractions showed the presence of purified protein, which was subsequently concentrated and assayed (Figure 2.3). Cleavage of the orange all-*trans*-*β*-apo-8'-carotenal by LeCCD1a results in the formation of a light yellow / green colour as the substrate is consumed (Figure 2.4). Initial assays with purified LeCCD1a (≈100 ng) in the presence of all-*trans*-*β*-apo-8'-carotenal (3 µg), however, showed neither a decrease in absorbance nor visual colour change versus time, indicating a loss of activity within the purified protein. Upon repeating, some samples showed only a minor decrease in absorbance at 485 nm, along with a slight decolouration of the reaction mixture upon visual inspection. However, the level of activity shown by these

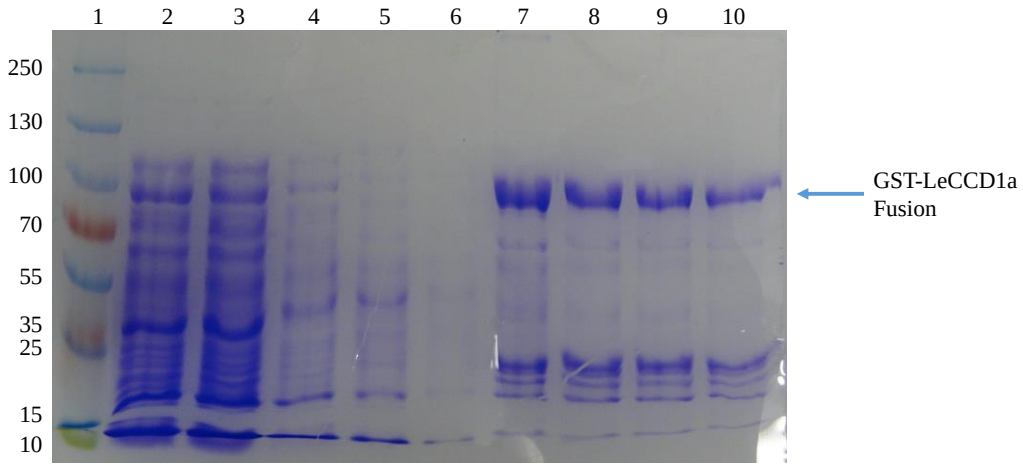


Figure 2.3 – 8% SDS-PAGE gel of fractions from the purification of glutathione tagged LeCCD1a. 1 – Ladder; 2 – Total soluble protein; 3 – Flow through; 4 – Wash fraction 1; 5 – Wash fraction 10; 6 – Wash fraction 20; 7 – Elution fraction 1; 8 – Elution fraction 2; 9 – Elution fraction 3; 10 – Elution fraction 4. Molecular weights of ladder are in kDa. Molecular weight for GST-LeCCD1a: 87.86 kDa (calculated from sequence).

purified samples, 0.36 µM min^{-1} was both lower than the cell lysate and significantly lower than would be expected for a purified sample. Control experiments with cell lysate containing the LeCCD1a protein overexpressed (≈50 µg total protein) showed an activity of 8.9 µM min^{-1}, with a visual colour change and decrease in absorbance (Figure 2.4). Additional controls with boiled enzyme and assays in the absence of the enzyme all failed to produce a decrease in absorbance at 485m. These results suggested that enzyme activity was lost during the purification process.

The purification of non-heme iron dependent enzymes has been reported whereby the iron cofactor is first removed from the enzyme in the cell lysate (Bugg 1993). The apo-enzyme is then purified before the iron is added back into the enzyme via an iron salt, restoring enzyme activity. By removing the catalytic iron centre, the enzyme can be purified before the protein is deactivated. Hence several changes were made to the purification technique. Firstly, 10% (v/v) glycerol, which is known to aid the stability of dioxygenases, and 0.1% Triton X-100 were added to the assay buffer to aid protein stability and prevent the aggregation of the protein due to potential hydrophobic patches (Bugg 2003). The reducing agent 2-mercaptoethanol was added to ensure that any iron present was kept in the iron (II) state. Iron

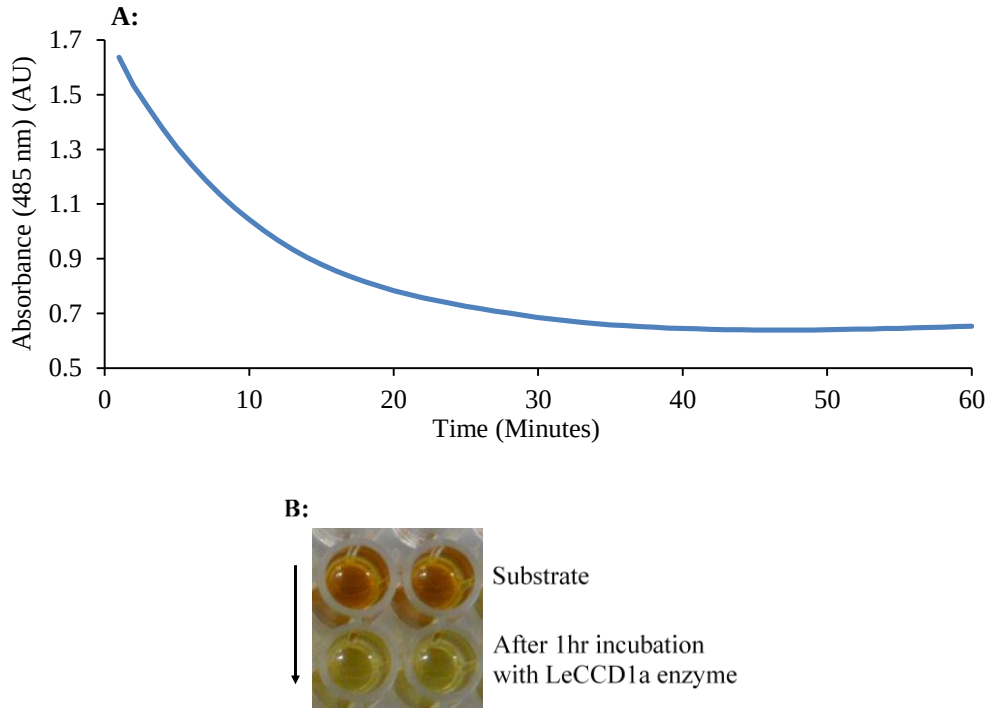


Figure 2.4 – A: Absorbance change at 485 nm shown after incubation of all-*trans*-*β*-apo-8'-carotenal with cell free extract containing overproduced LeCCD1a enzyme. **B:** Colour change observed on incubation of LeCCD1a with all-*trans*-*β*-apo-8'-carotenal.

was removed via the addition of 10mM 1, 10-phenanthroline to the cell lysate prior to purification, chelating the iron from the active site. Addition of 1, 10-phenanthroline was accompanied by the appearance of a red colour in the lysate, indicating iron chelation. It was hoped that removal of the iron would allow the apoenzyme to be stored for long periods at -80º C without loss of activity. LeCCD1a was purified as before and elution fractions containing the enzyme, as determined by SDS-PAGE gel electrophoresis, were pooled and concentrated. Purified LeCCD1a was then reactivated by the addition of 1 mM iron sulfate and 1mM BME. Enzymatic assay with purified enzyme showed activity characteristic of the LeCCD1a enzyme (Figure 2.5).

By comparison, the activity of the purified enzyme, 3.9 µM min^{-1}, appears to be lower than that of the cell lysate. This is most likely due to incomplete reactivation of the enzyme. It is also possible that the iron cofactor is required for maintaining the structure of the protein. In the absence of this cofactor, a percentage of the protein in the sample could have denatured,

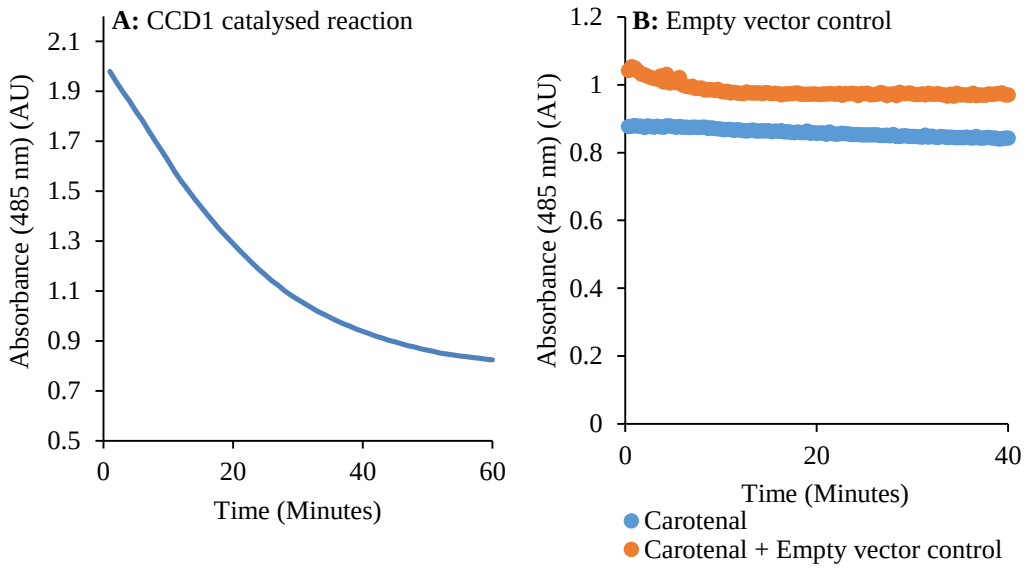


Figure 2.5 – A: Absorbance change at 485 nm shown after incubation of all-*trans*-*β*-apo-8'-carotenal with purified apo-LeCCD1a enzyme reconstituted with iron (II) sulfate. **B:** Control reaction with *E. coli* cell lysate with empty pGEX-4T-1 vector.

leading to the decreased activity. This method of purification has been subsequently used for the purification of *S. lycopersicum* LeNCED1 and *Z. mays* NCEDs as detailed in Chapter 3.

2.3 Structure-Activity Relationship of Hydroxamic Acid CCD Inhibitors towards *S. lycopersicum* CCD1a

Novel hydroxamic acid based inhibitors of carotenoid cleavage dioxygenases were first reported by Sergeant *et al*. in 2009. However, since the publication of these compounds, subsequent inhibitors have been developed, and a complete structure-activity profile has not been performed. Inhibition data of the hydroxamic acid and abamine CCD inhibitors against LeCCD1a is shown in Tables 2.1-2.5 (some data reproduced from Sergeant *et al*. 2009, marked with '*', Sergeant *et al*. 2013, marked with '‡' and Van Norman *et al*. 2014, marked with '!'). Representative data from an inhibition assay is shown in Figure 2.6. In addition to the hydroxamic acids, a selection of other compounds have also be assayed for inhibition against LeCCD1a. In particular, abamine compounds, first developed by Han *et al*. for the inhibition of *Vigna unguiculata* (cowpea) NCED (Han *et al*. 2004), and thiosemicarbazones (provided by Louisa Noffke and Prof. Peter Sadler, University of Warwick Chemistry, represented in Figure 2.7), which are known iron chelators used in medicine (Yu *et al*. 2009).

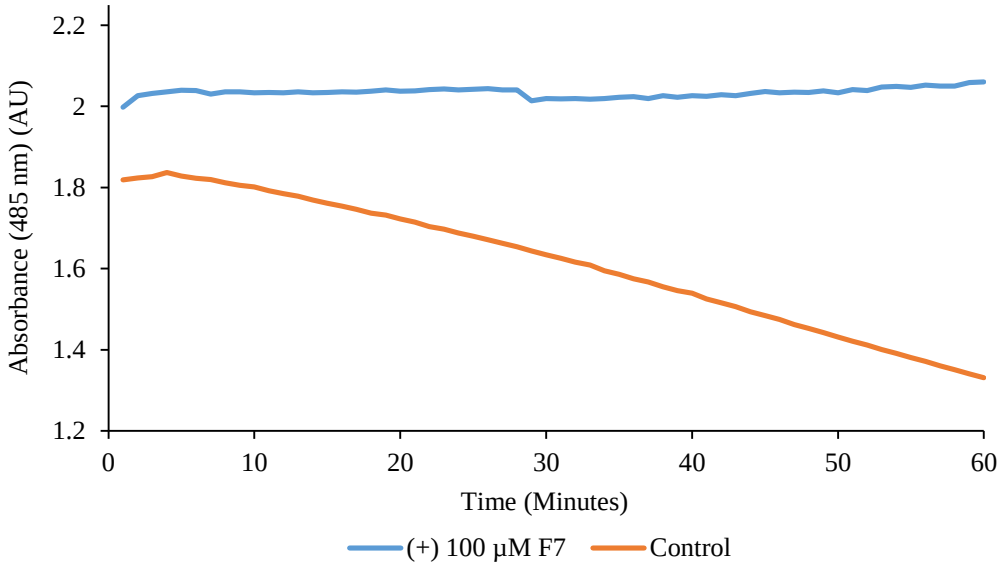


Figure 2.6 – Absorbance change at 485 nm for control cleavage reaction of LeCCD1a with all-*trans*-*β*-apo-8'-carotenal and for the same reaction in the presence of the inhibitor F7 at 100 µM. Control reaction is normal enzyme assay of LeCCD1a with all-*trans*-*β*-apo-8'-carotenal.

IC50 values reported are based on assays using cell free extract, therefore the total concentration of LeCCD1a is unknown. As such, exact comparison of data with other enzymes is not possible.

Most interestingly, the thiosemicarbazones show no inhibition towards CCD1. The thiosemicarbazones assayed contain phenyl groups and π systems similar to the hydroxamic acid and abamine inhibitors, bearing a resemblance in particular to D12 and D13, which showed 26% and 46% inhibition respectively against LeCCD1a at 100 µM. Thiosemicarbazones are known iron chelators, and as such, it was expected that these thiosemicarbazones would show activity against LeCCD1a (Yu *et al*. 2009). However, despite this, none of the compounds showed any inhibition at 100 µM. This suggests that the CCD enzymes are not susceptible to all metal chelators, and there is a specific element of the

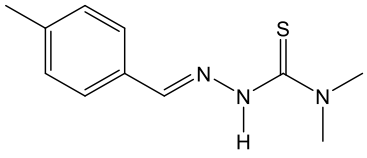


Figure 2.7 – Representative structure of thiosemicarbazones assayed.

hydroxamic acids which aids inhibition, or that the iron binding affinity of the thiosemicarbazones is lower than that of the hydroxamic acids. Hydroxamic acids have a pK_a around physiological pH (Green *et al*. 1958). As such, at assay conditions (pH 7.4) the *N*-alkoxy of the hydroxamic acid may well be deprotonated, resulting in a hydroxamic anion which could enhance binding to the metal with respect to the thiosemicarbazones (Figure 2.8). Oxidation of the iron (II) to iron (III) is then subsequently likely to enhance the binding of the hydroxamic to the iron (Sergeant *et al*. 2013) (see Chapter 3.5).

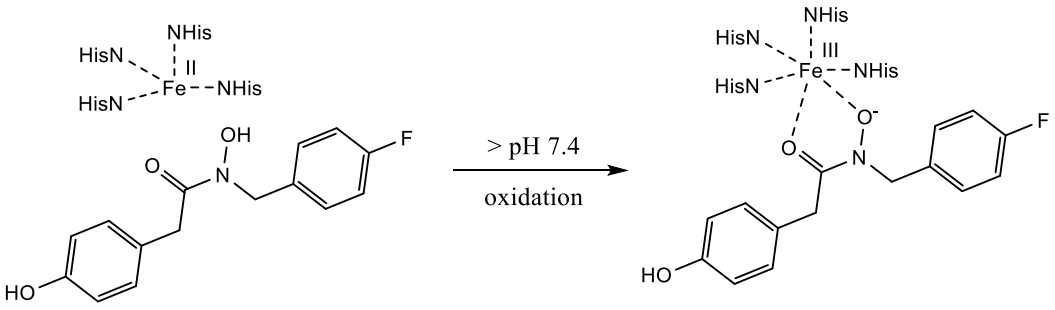


Figure 2.8 – Representation of deprotonation of hydroxamic acids at pH > 7.4, resulting in a hydroxamic acid anion which may aid binding to the active site iron (II).

The abamine based compounds assayed were all less effective than the hydroxamic acids. Abamine showed only 49% inhibition at 100 μM, comparable to the effect seen against *V. unguiculata* NCED. Compounds E1 to E4, based on the structure of abamine, vary in the length of the alkyl chain connecting the methyl ester to tertiary amine, which should increase the overall hydrophobicity of the inhibitor and potentially aid binding in the hydrophobic active site. However, increased chain length also increases the risk of steric clashes within the active site. In addition, the E series compounds have para substituted methoxy group on the phenyl ring as opposed to a dimethoxy substituted ring at the meta and para positions as in the case of abamine. The inhibitors E1 to E4 all showed 0% inhibition at 100 μM. The lack of inhibition shown by the E series compounds against LeCCD1a is quite surprising, given that abamine showed almost 50% inhibition. In abamine, the presence of two methoxy groups will increase the electron density in the phenyl group, due to the electron donating effects of methoxy groups, thus enhancing any potential π-π stacking interactions. Decreased inhibition

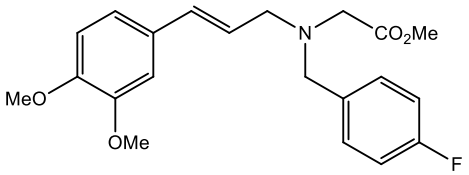
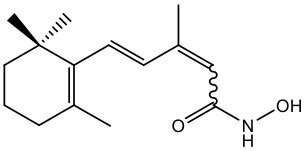
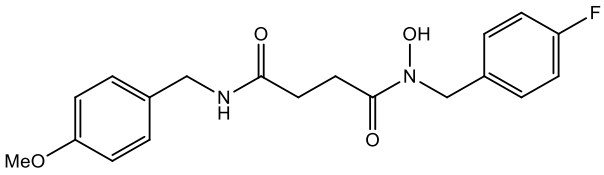
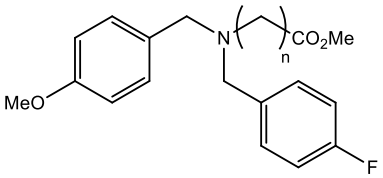
Inhibitor	Structure	Value 'n'	LeCCD1a Inhibition at 100 µM (%)	LeCCD1a IC_{50} (µM)
Abamine		--	49	210
B1*		--	90	20
B2		--	47	N/A
E1		1	0	N/A
E2		2	0	N/A
E3		3	0	N/A
E4		4	0	N/A

Table 2.1 – Inhibition data for hydroxamic acid compounds abamine, B1, B2 and E1-4 versus LeCCD1a. Compound marked with '*' previously published in Sergeant *et al*. 2009.

for the E series compounds would have been expected, given the fact that the decreased electron density in the phenyl group would have weakened π-π stacking interactions. However, some inhibition would still have been expected. This was not observed, and the reasons for the absence of any inhibition are unknown.

Without a doubt, the most effective inhibitors are D1 to D7. These compounds all show greater than 95% inhibition at 100 µM. D1, D2 and D3 have the lowest IC_{50} values at 0.9, 0.8 and 0.8 µM respectively. Hydroxyl groups on the terminal aromatic ring appear to increase the affinity for binding. The effect of these hydroxyl groups could be two-fold, through either a hydrogen bonding interaction within the active site, or increased donation of electron density into the phenyl ring, as discussed previously. Given the hydrophobicity of the carotenoid substrates, it is likely through increased electron density in the aromatic ring. In the case of D16 and D17, there is the added possibility that there is a steric effect brought

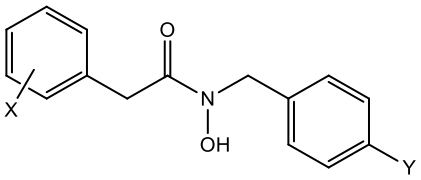
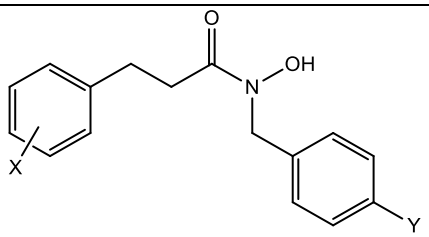
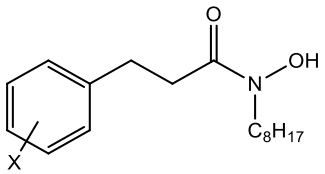
Inhibitor	Structure	Substituent X	Substituent Y	LeCCD1a Inhibition at 100 μM	LeCCD1a IC_{50}
D1*		4-Hydroxy	H	>95	0.9
D2*		4-Hydroxy	F	>95	0.8
D3*		3,4-Dihydroxy	F	>95	0.8
D4*		4-Methoxy	F	>95	2.5
D5*		3,4-Dimethoxy	H	>95	8.0
D6*		3,4-Dimethoxy	F	>95	9.0
D7*		4-Methoxy	F	>95	3.0
D16		4-*O*-Butyl	H	25	N/A
D17		4-Acetyl	H	24	N/A
D8*		3,4-Dimethoxy	H	61	N/A
D9*		4-Methoxy	H	>95	10.0
D14		3,4-Dimethoxy	F	N/A	N/A
D15[!]		4-Methoxy	F	70	10.0
D10*		3,4-Dimethoxy	---	65	N/A
D11*		4-Methoxy	---	53	N/A

Table 2.2 – Inhibition data for hydroxamic acid compounds D1-D11 and D14-D16 versus LeCCD1a. Compounds marked with '*' previously published in Sergeant *et al*. 2009. Compound marked with '!' previously published in Van Norman *et al*. 2014.

about by the bulky *O*-acetyl and *O*-butyl para substituents, causing a decrease in the inhibition seen.

Inhibitors D8, D14, and D15 all show less inhibition against LeCCD1a. These compounds have an increased alkyl chain length connecting the carbonyl of the hydroxamic acid to a phenyl group (ethylene linker as opposed to a methylene linker). This increased linker could be preventing the optimal alignment of the phenyl group for a π-π stacking interaction. The exception here is D9, which shows almost complete inhibition even with the ethylene linker. The IC_{50} of D9 (10 μM) is almost 10 times higher than D1-D3. Para fluorination of the phenyl group on the hydroxylamine has little to no effect.

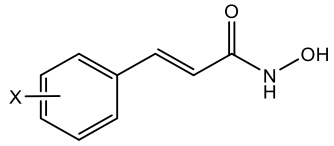
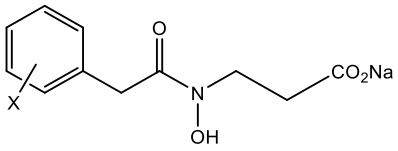
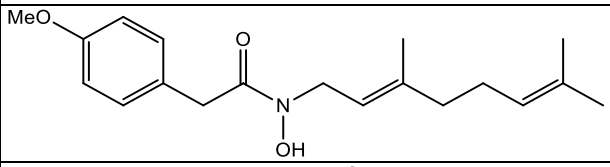
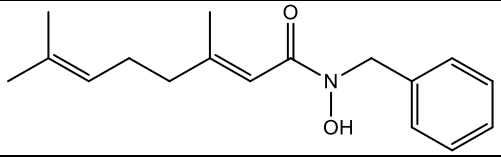
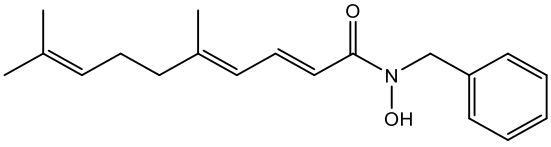
Inhibitor	Structure	Substituent X	LeCCD1a Inhibition at 100 µM	LeCCD1a IC_{50}
D12*		3,4-Dimethoxy	26	N/A
D13*		4-Methoxy	46	N/A
D20		4-Methoxy	0	N/A
D21		Napthyl	64	N/A
D30		---	>95	1.0
D31		---	90	15.0
D32		---	>95	1.0

Table 2.3 – Inhibition data for hydroxamic acid compounds D12-D13, D20-D21 and D30-D31 versus LeCCD1a. Compounds marked with '*' previously published in Sergeant *et al*. 2009.

Inhibitors D10-D13, D20 and D21 lack phenyl groups on the *N*-alkoxy terminal of the hydroxamic acid. D10 and D11 both have a saturated eight carbon chain, and show moderate inhibition of LeCCD1a at 100 µM, possibly due to a loss of π-π stacking interactions. Compounds D12 and D13 have a protonated *N*-alkoxy hydroxamic acid. Unsurprisingly, both D12 and D13 show poor inhibition against LeCCD1a, most likely due to a loss of hydrophobic interactions. The inhibitors D20 and D21 have a carboxylate on the *N*-alkoxy of the hydroxamic acid, but neither showed good inhibition. Given the pK_a of aliphatic carboxylic acids is around 5.0, it is probable that at the reaction pH (pH 7.4) the carboxylic acid will be deprotonated to a carboxylate anion. This anion is unlikely to bind well within an active site

designed for hydrophobic substrates, likely resulting in the lower levels of inhibition observed. It would be interesting to observe what effect, if any, replacing the carboxylate on D21 with benzyl moiety would have. One would expect that the increased π-π staking capabilities of a napthyl group would lead to an effective inhibitor.

Inhibitors D30, D31 and D32 all have long geranyl chains. These geranyl chains mimick the hydrophobic nature of the carotenoid chain and show greater than 90% inhibition at 100 µM. This inhibition data suggests that the positioning of the geranyl chain (either on the *N*-alkoxy or carboxylate of the hydroxamic acid) is irrelevant. This raises the question as to what the orientation of the inhibitors is within the active site. Do they bind with the *N*-alkoxy terminus first, or the carboxylate terminus first? D30 and D31 both have IC_{50} values of 1.0 µM, comparable to D1, D2 and D3. The 'H' series, H1-H4, are somewhat similar to the D30-D32, with long hydrophobic chains, no doubt aiding their binding within the active site. These compounds all show approximately 80% inhibition at 100 µM.

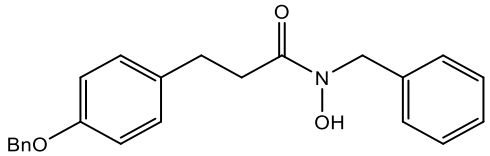
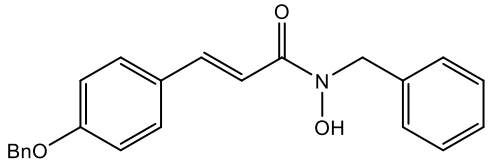
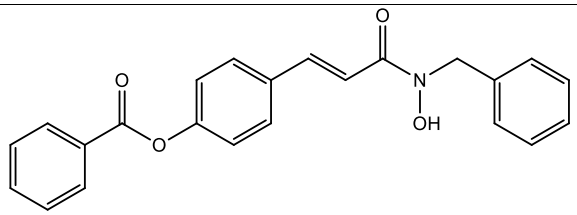
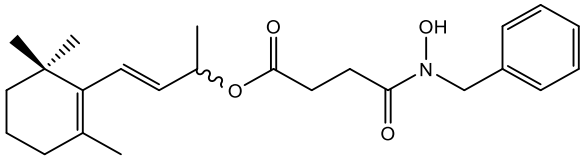
Inhibitor	Structure	LeCCD1a Inhibition at 100 µM	LeCCD1a IC_{50}
H1		>95	<10
H2		82	20.0
H3		>95	<10
H4		81	10.0

Table 2.4 – Inhibition data for hydroxamic acid compounds H1-H4 versus LeCCD1a. Bn represents a benzyl group (CH_2Ph).

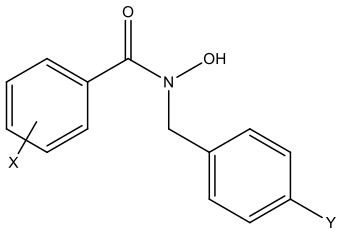
Inhibitor	Structure	Substituent X	Substituent Y	LeCCD1a Inhibition at 100 μM	LeCCD1a IC_{50}
F1*‡		4-Methoxy	H	>95	2
F2*‡		4-Methoxy	F	>95	2.5
F3*		3,4-Dimethoxy	H	50	N/A
F4*		3,4-Dimethoxy	F	0	N/A
F5‡		3-Chloro	H	0	N/A
F6‡		3-Amino	H	>95	<1
F7		3-Bromo	H	>95	<1
F8‡		3-Chloro-4-Methoxy	H	30	N/A

Table 2.5 – Inhibition data for hydroxamic acid compounds F1-F8 versus LeCCD1a. Compounds marked with '*' previously published in Sergeant *et al*. 2009. Compounds marked with an '‡' previously published in Sergeant *et al*. 2013.

Inhibitors F1, F2, F6 and F7 all also show greater than 95% inhibition at 100 μM, yet lack a connector between the carboxylate of the hydroxamic acid and the phenyl group. F6 and F7, which are substituted at the meta position, show comparable activity to F1 and F2, which are substituted at the para position. Inhibitors F3 and F4, differ from F1 and F2 in that they are additionally substituted at the meta position with an additional methoxy group. Both inhibitors gave less than 50% inhibition at 100 μM. Given how F1, F2, F6 and F7 all showed >95% inhibition with meta or para substitutions, it is intriguing why the dimethoxy substituted inhibitors would be less effective. This indicates that some sort of interaction at the three position is important. Interestingly though, a 3-amino or 3-chloro group increases activity heavily. Again, the presence of a 4-fluoro-phenyl group makes little difference. Inhibition data, in the form of percentage activity of enzyme in the presence of inhibitor is summarised in Figure 2.9.

Clearly there is a complex series of interactions within the active site that are required for optimal inhibitor binding. Additionally, several other matters complicate the analysis. For example, it is unknown in which orientation the inhibitors bind. Messing *et al*. showed that *Z.*

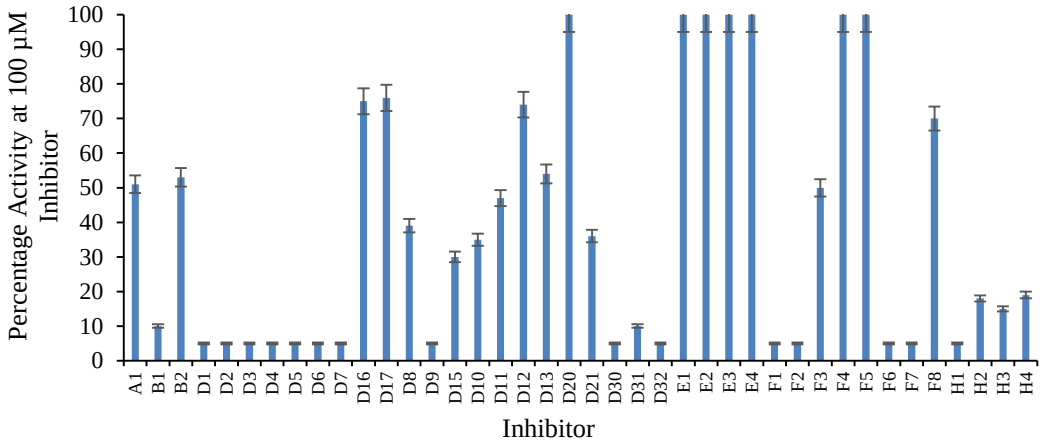


Figure 2.9 – Summary of inhibition data for hydroxamic acids against LeCCD1a.

mays Vp14 and *Z. mays* CCD1 from maize contained a high degree of sequence similarity (Messing *et al*. 2010). Using this data, they were able to model the structure of the ZmCCD1 enzyme. There is a high degree of sequence similarity between Vp14, ZmCCD1 (33% identity) and LeCCD1a (41% similarity). Using the active site model they produced (Figure 2.10, sequence alignment is shown in Figure 2.11), it is possible to interpret some of the data obtained from these inhibitors. The majority of the substrate binding tunnel consists of non-polar residues. However, three other key pieces of information can be taken from the active site. Firstly, the presence of at least two charged residues which are in close proximity to the metal centre, secondly, the fact that the active site tunnel appears to turn once it has passed the iron centre and thirdly, the presence of a small aperture before the metal centre.

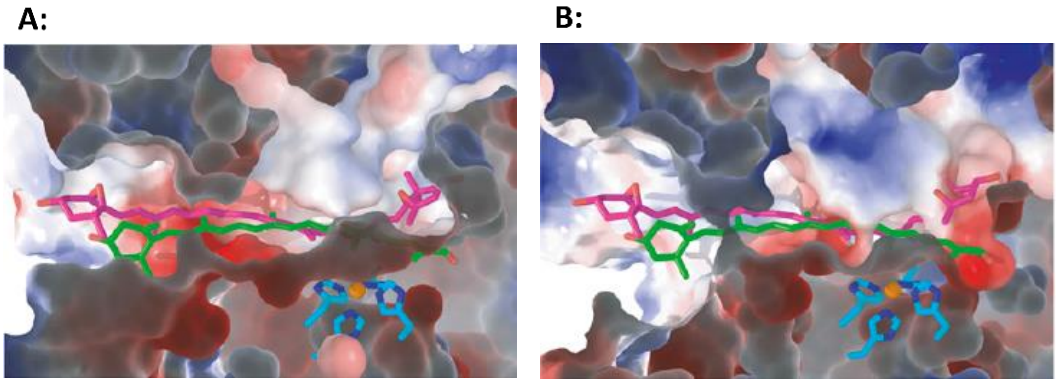


Figure 2.10 – Modelling of 9'-*cis*-neoxanthin (pink) and all-*trans*-*β*-apo-8'-carotenal (green) within the active site of ZmNCED1 (**A**) and ZmCCD1 (**B**). Figure taken from Messing *et al*. 2010.

```
ZmNCED1           MQGLAPPTSVSIHRHLPARSRARASNSVRFSPRAVSSVPPAECLQAPFHKPVADLPAPSR
ZmCCD1            ------------------------------------------------------MGTEAE
LeCCD1a           ------------------------------------------------------MG---R
                                                                           :    .
ZmNCED1           KPAAIAVPGHAAAPRKAEGGKKQLNLF--QRAAAAALDAFEEGFVANVLERPHGLPSTAD
ZmCCD1            QPD--------MDSHRNDGVVVVPAPRPRKGLASWALDLLESLAVRL-------GHDKTK
LeCCD1a           KED--------DGVERIEGGVVVVNPKPRRGITAKAIDLLEWGIVKL-------MHDSSK
                  :             .: :*          :  :: *:* :*   *          ..:.
ZmNCED1           PAVQIAGNFAPV-GERPPVHELPVSGRIPPFIDGVYARNGANPCFDPVAGHHLFDGDGMV
ZmCCD1            PLHWLSGNFAPVVEETPPAPNLTVRGHLPECLNGEFVRVGPNPKFAPVAGYHWFDGDGMI
LeCCD1a           PLHYLQGNFAPT-DETPPLNDLVVQGHLPECLNGEFVRVGPNPKFAPVAGYHWFDGDGMI
                  *   : *****.  * **  :* * *::*  ::* :.* * ** * ****:* ******:
ZmNCED1           HALRIRNGAAESYACRFTETARLRQERAIGRPVFPKAIGELHGHSGIARLALFYARAACG
ZmCCD1            HAMRIKDGKA-TYVSRYVKTARLKQEEYFGGAKFMK-IGDLKGFFGLFMVQMQQLRKKFK
LeCCD1a           HGLRIKDGKA-TYVSRYVRTSRLKQEEFFGGAKFMK-VGDLKGLFGLFTVYMQMLRTKLK
                  *.:**::* * :*..*:..*:**:**. :*   * * :*:*:*  *:  : :   *
ZmNCED1           LVDPSAGTGVANAGLVYFNGRLLAMSEDDLPYHVRVADDGDLETVGRYDFDGQLGCAMIA
ZmCCD1            VLDFTYGFGTANTALIYHHGKLMALSEADKPYVVKVLEDGDLQTLGLLDYDKRLKHSFTA
LeCCD1a           VLDISYGNSTANTALVYHHGKLLALSEADKPYALKVLEDGDLQTLGMLDYDKRLTHSFTA
                  ::* : * ..**:.*:*..*:*:*:** * ** ::* :****:*:*  *:* :*  :: *
ZmNCED1           HPKLDPATGELHALSYDVIKRPYLKYFYFRPDGTKSDDVEIPLEQPTMIHDFAITENLVV
ZmCCD1            HPKVDPFTDEMFTFGYSH-EPPYCTYRVINKEGAMLDPVPITIPESVMMHDFAITENYSI
LeCCD1a           HPKVDAVTGEMFTFGYAH-TPPYITYRVISKDGIMQDPVPITIPEPIMMHDFAITENYAI
                  ***:*  * *:.::.*     ** .*  :  :*   * * * : :  *:******** :
ZmNCED1           VPDHQVVFKLQEMLRGGSP-VVLDKEKTSRFGVLPKHAADASEMAWVDVPDCFCFHLWNA
ZmCCD1            FMDLPLLFRPKEMVKNGEFIYKFDPTKKGRFGILPRYAKDDKLIRWFQLPNCFIFHNANA
LeCCD1a           MMDLPLCFRPKEMVKKNKLAFTFDATKKARFGVLPRYANNEALIRWFELPNCFIFHNANA
                  . *  : *: :**::  .    :*  *..***:**::* :   : *.::*:** **  **
ZmNCED1           WEDEATGEVVVIGSCMTPA---DSI---FNESDERLESVLTEIRLDARTGRSTRRAVLPP
ZmCCD1            WEEGDE---VVLITCRLENPDLDKVNGYQSDKLENFGNELYEMRFNMKTGAASQKQLS--
LeCCD1a           WEEGDE---VVLITCRLVNPDLDMVNGAVKEKLENFCNELYEMRFNMKSGAASQKKLS--
                  **:      **: :*       * :    .:. *.: . * *:*:: ::* :::: :
ZmNCED1           SQQVNLEVGMVNRNLLGRETRYAYLAVAEPWPKVSGFAKVDLSTG------------EL
ZmCCD1            --VSAVDFPRVNESYTGRKQRYVYCTILDSIAKVTGIIKFDLHAEPESGVKVLEVGGNVQ
LeCCD1a           --ESAVDFPRINENYTGRKQRYVYGTTLNSIAKVTGVIKFDLHAEPETGKSQLEVGGNVQ
                       ::.  :*..  **: **.* :  :   **:*. *.** :
ZmNCED1           TKFEYGEGRFGGEPCFVPMDPAAAHPRGEDDGYVLTFVHDERAGTSELLVVNAADMRLE-
ZmCCD1            GIYDLGPGRFGSEAIFVPKHPGV--SGEEDDGYLIFFVHDENTGKSEVNVIDAKTMSADP
LeCCD1a           GIFDLGPGRFGSEAVFVPSRPGT--EREEDDGYLIFFVHDENTGKSAVNVIDAKTMSAEP
                    :: * ****.*  ***  *..     *****:: *****.:*.* : *::*  *  :
ZmNCED1           -ATVQLPSRVPFGFHGTFITGQELEAQAA--
ZmCCD1            VAVVELPNRVPYGFHAFFVTEDQLARQAEGQ
LeCCD1a           VAVVELPKRVPYGFHAFFVTEEQIQEQAKL-
                   *.*:**.***:***. *:* :::  **
```

Figure 2.11 – Sequence alignment of ZmNCED1, ZmCCD1 and LeCCD1a using ClustalΩ (Goujon *et al*. 2010, Sievers *et al*. 2011). Accession numbers: ZmNCED1 - O24592; ZmCCD1 - Q45VT7; LeCCD1a - Q6E4P5. Conserved histidine residues are shown in red.

The tertiary amines abamine and E1-E4 are all sterically bulky. The active site of LeCCD1a is likely to be linear due to the all-*trans* nature of the carotenoid substrate, restricting the binding of the abamines into the active site. It is possible that the hydroxyl groups of D1, D2, D3 and D4 are interacting with the positively charged groups located near the active site. The fact that D4-D8 all show higher IC_{50} values than D1-D3 seems to confirm the presence of some sort of polar-charged interaction, since ethers are less polar than alcohols.

The increased chain length of D8, D9, D14 and D15 may position the aromatic ring in such a position that it interferes with the aperture close to the active site. This would result in the decreased binding affinities seen. The increased chain length on D30, D31 and D32 potentially increases the hydrophobic interaction within the active site, resulting in their high levels of inhibition. This would also explain the respectable levels of inhibition reported for the 'H' series.

Although the 'F' series compounds F1 and F2 both lack a hydroxyl group, it is possible that the absence of an alkyl chain connecting the C-terminus of the hydroxamic acid with the phenyl group consolidates a hydrophobic interaction, maximising binding. Halogen-oxygen interactions in biological systems are known (Auffinger *et al*. 2004). As such, it appears likely that the excellent data observed for F5 and F6 inhibiting LeCCD1a is due to these moieties interacting favourable with one of the polar residues within the active site. However, F5 and F8, both with meta halogen substitutions would have been expected to behave in a similar way. It is possible that the meta methoxy groups are too large for the three position on the phenyl group in F3 and F4, sterically clashing with residues of the active site. The lower activity seen by D20 and D21 is somewhat surprising, given that the carboxyl group might have been expected to interact with one of the polar residues within the active site.

Without a definitive crystal structure of either ZmCCD1, or more importantly LeCCD1a, it is very difficult to assign accurate structure activity relationships based on the data here. A crystal structure with bound substrate or inhibitor would also help to explain possible discrepancies in the data observed.

2.4 Conclusions

It has been possible to purify the CCD1a enzyme from *S. lycopersicum* in the absence of the catalytic metal cofactor, as indicated by SDS-PAGE analysis. However, protein mass spectrometry would be required to definitively illustrate the purification of the native and apo-proteins. Reactivated apo-LeCCD1a has been shown to be able to retain catalytic activity.

From the structure activity data obtained for the LeCCD1a enzyme it is possible to elucidate which features of the hydroxamic acids are required for optimal activity against LeCCD1a. The key features of the inhibition assays are summarised in Figure 2.12. Hydroxamic acids were more potent inhibitors of LeCCD1a mediated carotenoid cleavage than abamines and thiosemicarbazones. One of the key features observed is that an element of hydrophobicity is required in the inhibitors, which is unsurprising given the hydrophobic nature of the substrates the enzyme utilises. Inhibitors which maximise this hydrophobicity, but are not sterically bulky are better inhibitors. Truncated hydroxamic acids with no groups on the *N*-alkoxy of the hydroxamic acid and carboxylate groups reduced inhibition, as did large bulky substitutions on the phenyl group of the carboxylate terminal of the hydroxamic acids. Hydroxyl and methoxy substitutions also lead to higher levels of inhibition, likely through promoting π-π stacking interactions in the active site. *In vitro*, no difference in inhibition was observed for the fluorine substituted inhibitors.

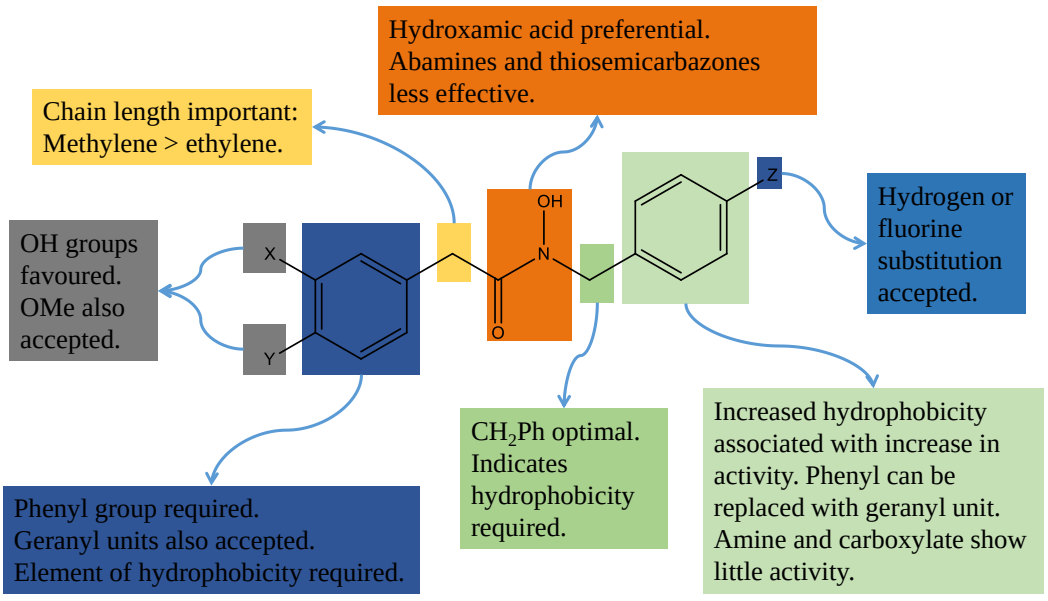


Figure 2.12 – Key features of hydroxamic acid compounds required for the inhibition of LeCCD1a.

2.5 Future Work

It would be of great interest to perform mechanistic studies on the CCD1 enzyme. Since CCD1 is a cytosolic enzyme, it is more amenable to spectroscopic studies. As a result, it would be interesting to investigate whether the enzymatic mechanism proceeds through a

dioxetane intermediate or through a Criegee rearrangement. In addition, there is no mechanistic proof that the CCD enzymes proceed through a carbocation intermediate. It would be possible to use a series of substituted stilbene derivatives to probe the existence of this intermediate, since different substitutions would have either electron donating or electron withdrawing effects, which would influence the stability of the carbocation and hence the rate of the cleavage reaction (Figure 2.13). Given that CCD1 is a relatively promiscuous enzyme, it would be hoped that stilbenes and stilbene derivatives would be accepted by the enzyme.

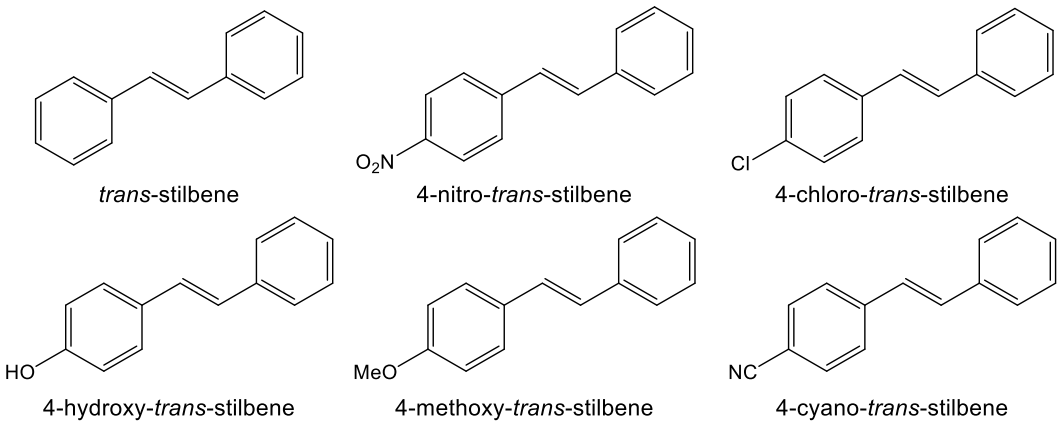


Figure 2.13 – Structures of trans-stilbene derivatives which could be used to probe the mechanism of LeCCD1a.

Iron is a common cofactor in oxygenase and dioxygenase enzymes (Bugg 2003). The consecutive and readily accessible oxidation states of iron make it ideal for the activation of molecular oxygen to superoxide. Other transition metals however, are also able to perform such a task. Manganese dependent oxygenases, for example, use a manganese (II) cofactor to activate molecular oxygen (Gibello *et al*. 1994, Whiting *et al*. 1996, Que *et al*. 1981) whilst quercetin 2,3-dioxygenase is the only known copper (II) dependent dioxygenases (Oka & Simpson 1971, Hund *et al*. 1999). Recently, Fielding and co-workers demonstrated that in the homoprotocatechuate 2,3-dioxygenase from *Brevibacterium fuscum* the catalytic iron (II) cofactor can be replaced with either manganese (II) or cobalt (II) with retention of activity (Fielding *et al*. 2011). In fact, in the presence of the cobalt (II) cofactor, the catalytic activity of the enzyme is enhanced and it would be interesting to see whether CCD1 could utilise other transition metal elements in its active site. Uptake of metals into the active site may be a

problem as such an alternative method for producing CCD1 enzyme substituted with other metals would be required, such as growing *E. coli* in minimal media, in the absence of iron. Adding in other metal salts, such as manganese (II) sulfate, when overproducing CCD1 would ideally force the incorporation of other metal ions into the active site. The incorporation of these metals could then be confirmed by Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES) or by ultraviolet-visible spectroscopy by the addition of a metal chelator.

Finally, as with all the CCD enzymes, the development of an inhibitor which is selective for solely the CCD1 enzyme would be desirable. A selective inhibitor for CCD1 would allow the roles of CCD1 *in planta* to be observed through phenotypic effects on plants. The inhibitors developed and assayed here could be used to probe the roles of CCDs in other organisms, such as in bacteria and mammals.

Chapter Three: Expression, Inhibition and Assays of 9'-*cis*-Epoxycarotenoid Cleavage Dioxygenases (NCED) from *Solanum lycopersicum* and *Zea mays*

3.1 Introduction

The enzyme 9'-*cis*-epoxycarotenoid cleavage dioxygenase 1 (NCED1, Vp14) from *Zea mays* (maize) was the first biochemically characterised carotenoid cleavage enzyme (Schwartz *et al.* 1997). Homologues of NCED have been found in a wide variety of different plant species such as *Arabidopsis thaliana* and *Solanum lycopersicum* (Tan *et al*. 2003, Burbidge *et al*. 1997). There are varying numbers of NCED homologues in different species. In *S. lycopersicum* there are only two NCEDs. However, in *Z. mays* and *A. thaliana* there are five different NCED homologues (Tan *et al*. 2003, Lefebvre *et al*. 2006).

The rate limiting step in the biosynthesis of the plant signalling hormone abscisic acid is catalysed by 9'-*cis*-epoxycarotenoid cleavage dioxygenases, which catalyse the selective oxidative cleavage of the 11',12' double bond in 9'-*cis*-neoxanthin or 9'-*cis*-violoxanthin to give a C_{25} aldehyde product and a second aldehyde called xanthoxin (Figure 3.1) (Taylor *et al*. 2005). NCEDs are selective solely for the 9'-*cis* isomer of violoxanthin and neoxanthin and show no activity against all-*trans* substrates (Schwartz *et al*. 2003). Assays of NCED are stopped and low throughput, requiring quantification of products by HPLC. Nevertheless, with these methods, a K_M for NCED1 from *Z. mays* has been reported of 27 μM for 9'-*cis*-neoxanthin and 58 μM for 9'-*cis*-violaxanthin (Schwartz *et al*. 2003).

The work within this chapter describes the use of low throughput assays for performing structure activity analysis of hydroxamic acid inhibitors against *S. lycopersicum* NCED1 and *Z. mays* NCED homologues. In addition development of high throughput assays for the screening of compounds against *Z. mays* NCED1 is described. Coupled assay work

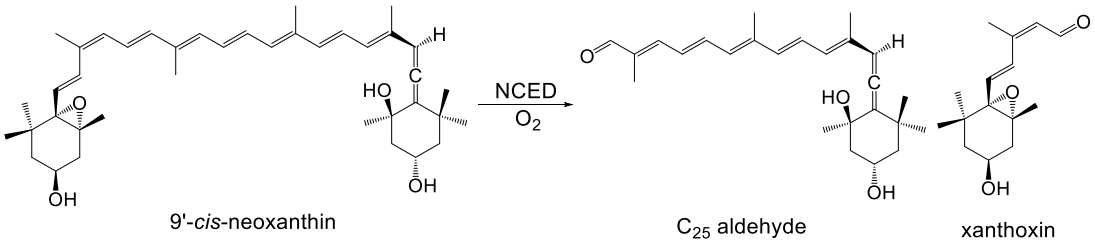


Figure 3.1 – Schematic representation of the NCED catalysed 11',12' cleavage of 9'-*cis*-neoxanthin to xanthoxin and C_{25} aldehyde.

was performed in partial collaboration with Dr David Brocklehurst (Syngenta, Jealott's Hill) and undertaken partially at Syngenta Jealott's Hill.

3.2 Expression and HPLC Assays of *S. lycopersicum* NCED1

The natural substrate for NCEDs, 9'-*cis*-neoxanthin is unstable, and susceptible to light, pH, temperature and oxygen. This high degree of sensitivity is due to the functional groups within the enzyme, such as the epoxide, allene and conjugated double bonds. 9'-*cis*-neoxanthin is available commercially, but the cost was too high for the planned study. As a result 9'-*cis*-neoxanthin was extracted from spinach leaves, which have a high natural abundance of neoxanthin (Tonucci *et al.* 1995). 9'-*cis*-neoxanthin was extracted from fresh spinach using previously reported methods (Sergeant *et al*. 2009). Initial attempts to characterise the 9'-*cis*-neoxanthin via thin layered chromatography (TLC), as published, proved inconclusive (Schwartz *et al*. 1997). However, the absorbance maxima in the ultraviolet-visible (UV-Vis) spectra of 9'-*cis*-neoxanthin and all-*trans*-neoxanthin have been reported (Table 3.1) (Britton 1985, Cholnokey *et al*. 1969). As such, it was possible to characterise the 9'-*cis*-neoxanthin via its characteristic UV-Vis spectra. The purified 9'-*cis*-

	All-*trans*-neoxanthin (Literature)	9'-*cis*-neoxanthin (Literature)	Purified 9'-*cis*-neoxanthin	Authentic 9'-*cis*-neoxanthin (CaroteNature)
Absorbance maxima (nm)	467	467	465	466
	440	439	439	437
	415	416	417	414

Table 3.1 - Absorbance maxima recorded for purified 9'-*cis*-neoxanthin compared with literature values for all-*trans*-neoxanthin and 9'-*cis*-neoxanthin (Britton 1985, Cholnokey *et al.* 1969).

neoxanthin was found to absorb at 417, 439 and 465 nm (Figure 3.2). Authentic 9'-*cis*-neoxanthin from CaroteNature (Switzerland, provided by Dr David Brocklehurst, Syngenta, Jealott's Hill) was analysed by UV-Vis spectroscopy. The spectra of the authentic sample matched with that of the 9'-*cis*-neoxanthin purified from spinach, with λ_{MAX} values of 414, 437 and 466 nm (Figure 3.2). Additionally, comparison of authentic 9'-*cis*-neoxanthin and purified 9'-*cis*-neoxanthin by HPLC showed that both had identical retention times (Figure 3.3). Using Beer-Lambert's law and authentic 9'-*cis*-neoxanthin, a molar extinction coefficient of 1.33 x 10^5 M^{-1} cm^{-1} was calculated for 9'-*cis*-neoxanthin (Beer 1852). The slight discrepancies between the literature values reported for 9'-*cis*-neoxanthin are likely due to *cis-trans* isomerisation of other alkenes on the polyene backbone or solvent effects. It was found that fresh spinach was required for assays. Spinach more than one day old gave a low yield of 9'-*cis*-neoxanthin and was not turned over by the enzyme.

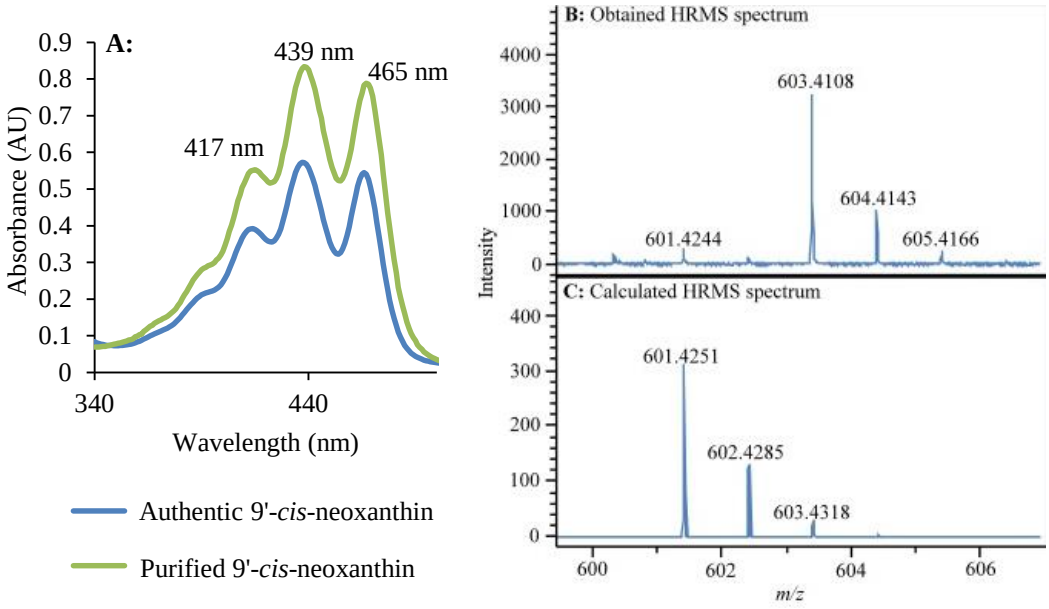


Figure 3.2 – A: Absorbance spectra of purified and authentic 9'-*cis*-neoxanthin. **B:** Obtained HRMS of purified 9'-*cis*-neoxanthin. **C:** Predicted HRMS spectrum. Peaks at *m/z* 603.4108, 604.4143 and 605.4166 in (**B**) are due to polyethylene glycol contamination.

Investigations into LeNCED1 involved the use of pGEX-4T-1-LeNCED1 pRosetta BL21 *E. coli* (GST-LeNCED1) provided by Dr Andrew Thompson (Cranfield) (accession number O24023). Cell free extract containing overproduced glutathione-S-transferase (GST) tagged LeNCED1 was prepared according to the method of Sergeant *et al.* 2009 and a band

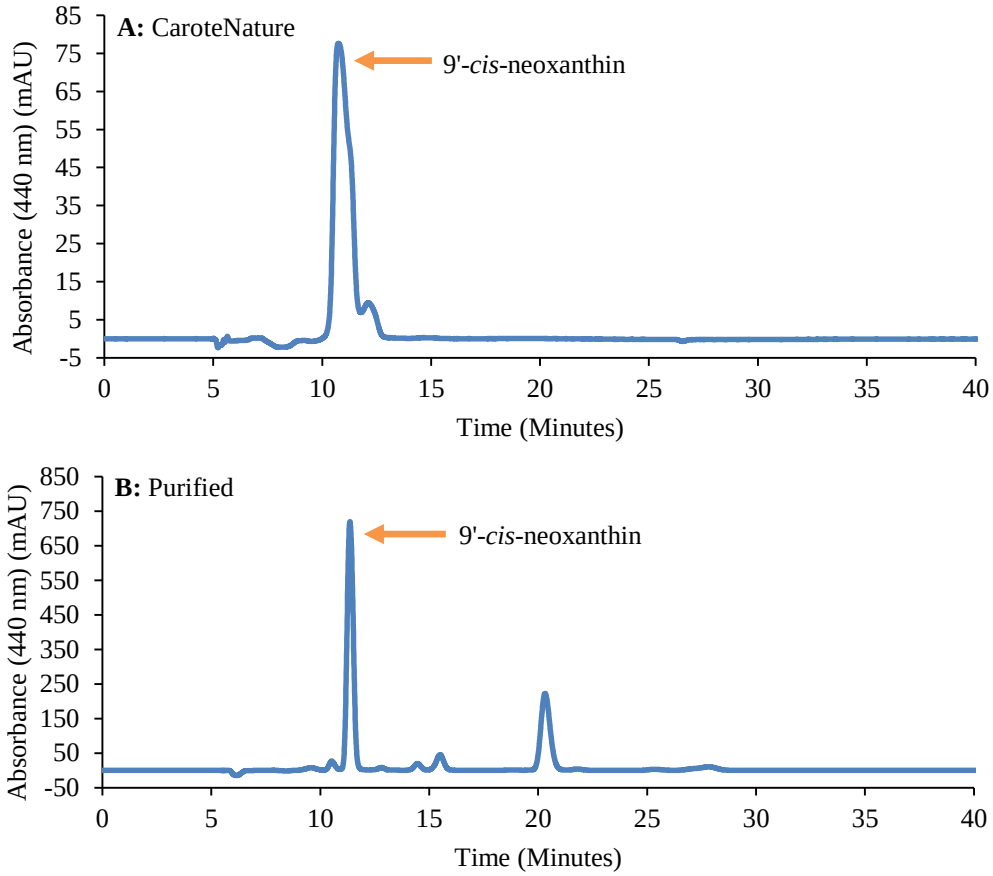


Figure 3.3 – HPLC chromatogram of authentic 9'-*cis*-neoxanthin (CaroteNature, **A**) and 9'-*cis*-neoxanthin purified from spinach (**B**). 9'-*cis*-neoxanthin is the main peak in both chromatographs.

of the correct molecular weight was visible via sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 3.4) (Sergeant *et al*. 2009). As with the LeCCD1a enzyme, cell free extract was used in preference to purified enzyme due to the instability of the enzyme.

Given that cell free extract containing overproduced LeNCED1 was used for the assays, it was not possible to determine the exact concentration of LeNCED1 present in the assays. However, approximately 100 µg of total *E. coli* lysate was added to each assay. A 2:1 (v/v) excess of cell lysate (from 3 mL of cell free extract prepared from a 500 mL culture of *E. coli*) containing overproduced GST-LeNCED1 to 9'-*cis*-neoxanthin was required to allow the reaction to proceed to 50% completion. Each assay contained approximately 60 µM of 9'-*cis*-neoxanthin (purified from spinach). Upon incubation of LeNCED1 with 9'-*cis*-neoxanthin

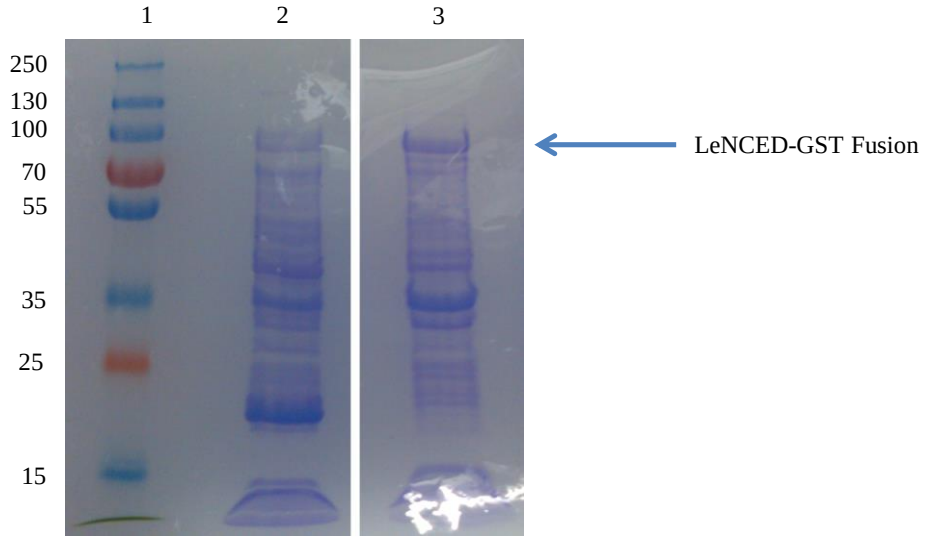


Figure 3.4 – 8% SDS-PAGE gel showing overproduced GST-LeNCED1 fusion in the cell free extract from pRosetta pGEX-LeNCED1 BL21 *E. coli*. 1 – Ladder; 2 – Non-induced control (NIC); 3 – Cell free extract from *E. coli* induced with IPTG. Molecular weights of ladder are in kDa. Molecular weight of GST-tagged LeNCED1: 93.9 kDa (calculated from sequence). Mass spectrometry analysis would be required to definitively confirm the presence of LeNCED1.

for 20 minutes at 25° C, a product peak at 6.5 minutes was observed at 440 nm, corresponding to that reported by Sergeant *et al*. 2009 (Figure 3.5A). Integration of the product and substrate peaks observed in the HPLC chromatogram was used to follow the course of the reaction.

A variety of other different strains encoding the LeNCED1 enzyme were obtained from Dr Andrew Thompson (Cranfield University) (Table 3.2). All constructs gave similar or

Construct	Enzyme Expressed	Features of strain
pGEX-4T-1-LeNCED1 BL21	GST-LeNCED1	Normal BL21
pGEX-4T-1-LeNCED1 pRosetta BL21	GST-LeNCED1	Contains Rosetta plasmid coding for expression of rare codons
p14b-LeNCED1 BL21	N-His_6-LeNCED1	Normal BL21
pGEX-4T-1-LeNCED1 Origami	GST-LeNCED1	*E. coli* K-12 strain lacking thioredoxin reductase and glutathione reductase to enhance disufide bond formation.
pGEX-4T-1-LeNCED1 Gami	GST-LeNCED1	Origami host strain containing the pRosetta plasmid

Table 3.2 – Comparison of the *E. coli* strains expressing LeNCED1 tested and the features provided by each strain (Bessette *et al*. 1999, Kane 1995, Merck Millipore Ltd 2012).

lower expression, and as such, all subsequent experiments were conducted with the pGEX-LeNCED1 pRosetta BL21 construct.

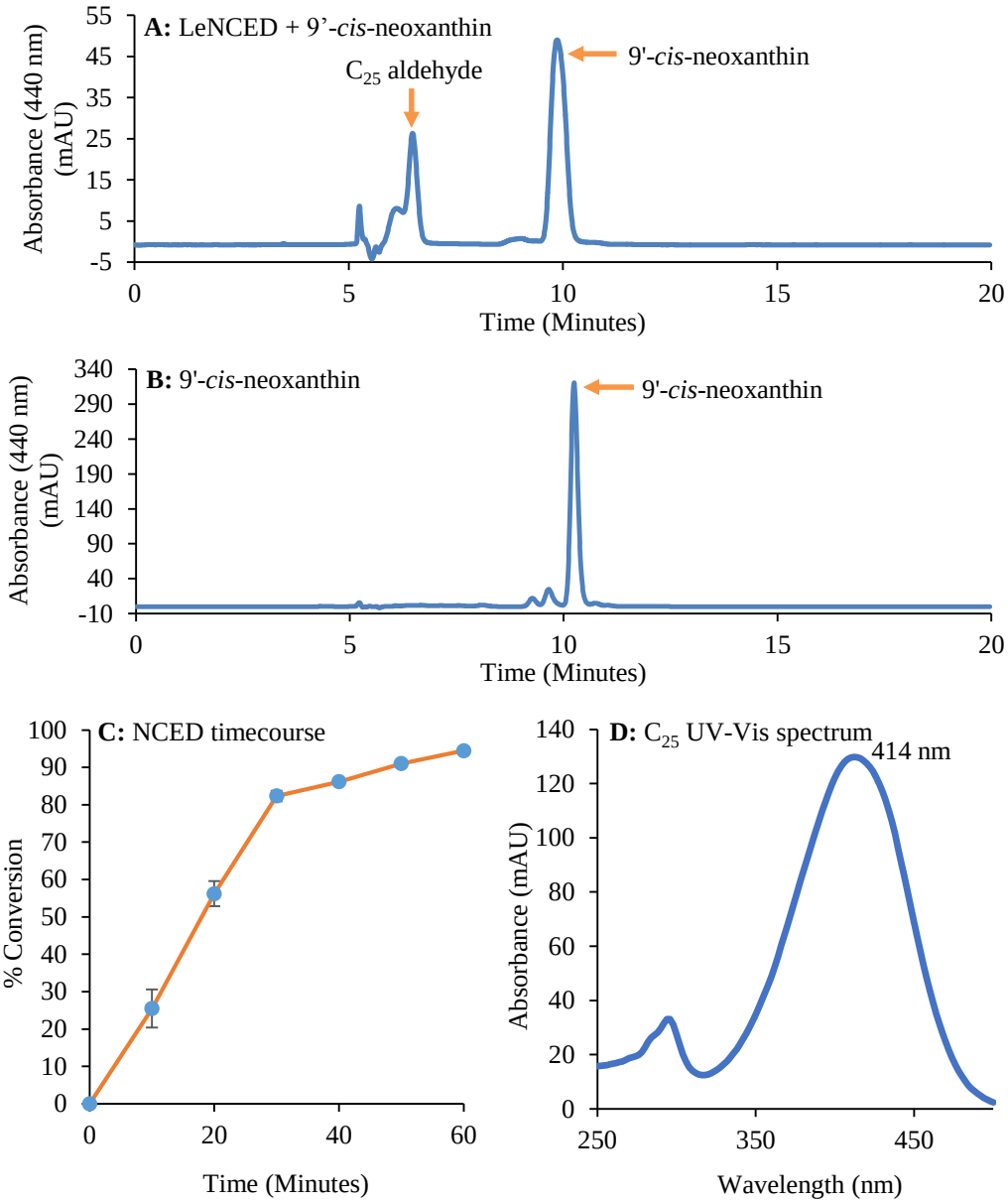


Figure 3.5 – HPLC chromatogram of LeNCED1 catalysed cleavage of 9'-*cis*-neoxanthin (**A**) and 9'-*cis*-neoxanthin (**B**). C_{25} aldehyde peak can be seen eluting at 6.5 minutes in the chromatogram of the enzyme catalysed reaction, consistent with the observations of Sergeant *et al*. 2009. (**C**) Time dependent increase in the magnitude of the peak at 6.2 minutes is consistent with an enzyme catalysed reaction. (**D**) UV absorption spectrum of the peak at 6.2 minutes gives a λ_{MAX} of 414 nm, similar to the theoretical value of 411 nm for the C_{25} aldehyde (Williams & Fleming 1995). High resolution mass spectrometry would be required for a definitive assignment of the peak at 6.2 minutes.

3.3 Structure Activity Relationship of Hydroxamic Acid CCD Inhibitors Towards *S. lycopersicum* NCED

All assays with LeNCED1 were conducted in a total volume of 200 µL with inhibitors at 100 µM. Percentage inhibition was calculated by working out the percentage conversion for each reaction from integration of the substrate and the product peaks for each assay at 440 nm. An example chromatogram comparing the inhibited and un-inhibited assays is shown in Figure 3.6.

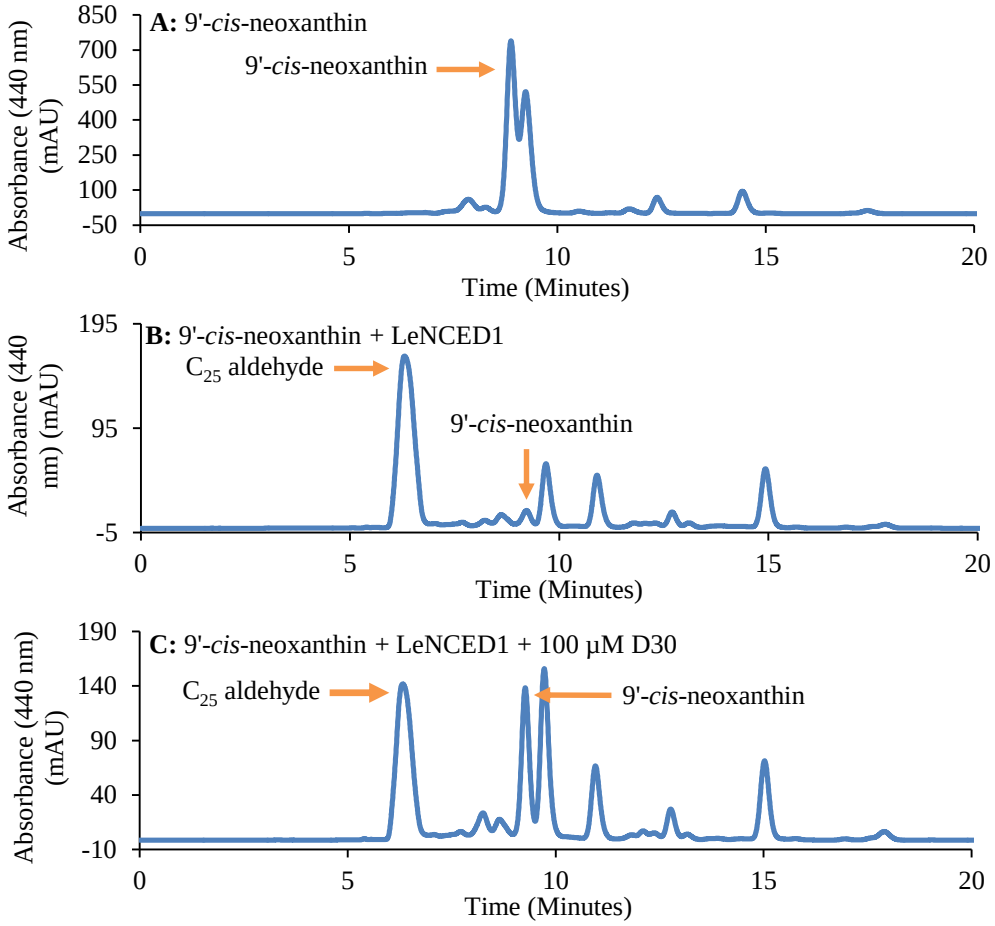


Figure 3.6 – HPLC chromatogram of 9'-*cis*-neoxanthin control (top, **A**), LeNCED1 catalysed cleavage of 9'-*cis*-neoxanthin in the absence (middle, **B**, 71% conversion) and presence (bottom, **C**, 50% conversion) of 100 µM D30. Additional peaks in chromatograph are believed to be other carotenoids co-purified from spinach.

Inhibition data of the hydroxamic acids and abamine CCD inhibitors against LeNCED1 is shown in tables 3.3-3.7 (some data reproduced from Sergeant *et al*. 2009, marked with '*', and Van Norman *et al*. 2014, marked with '!'). Although some inhibition was observed, neither the hydroxamic acids nor the abamines are effective inhibitors against

LeNCED1. Within the literature, three classes of compound have previously been shown to be inhibitors of NCED; nordihydroguaiaretic acid (NDGA), the tertiary amines abamine and abamineSG and the sesquiterpene based compounds (Han *et al*. 2004, N. Kitahata *et al*. 2006, Boyd *et al*. 2009 respectively). Against *Vigna unguiculata* (cowpea) NCED, abamine showed approximately 50% inhibition at 100 µM and abamineSG showed 80% inhibition of the same enzyme at 100 µM. However, against LeNCED1 abamine showed only 20% inhibition (Sergeant *et al*. 2009). Derivatives of abamine (compounds E1-E4) showed levels of inhibition similar to abamine itself. It is not clear why the levels of inhibition for the same compound are so different for homologues of the same enzyme, which have 70.1% identity.

Within the compounds of the D series it is difficult to observe any trends within the data. Compounds D1-D7, D16 and D17 are all structurally similar, with a methylene bridge

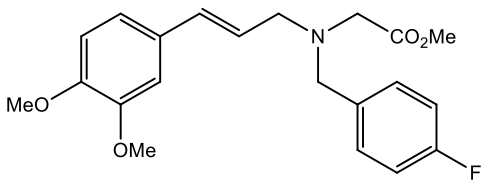
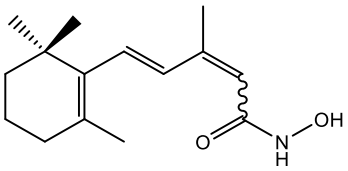
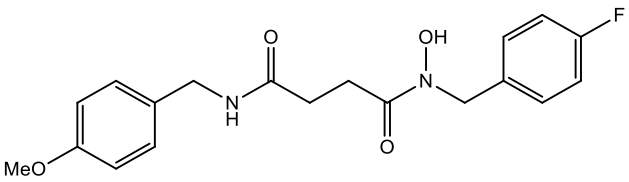
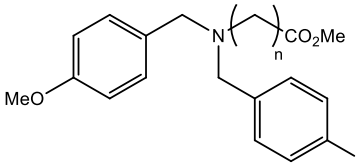
Inhibitor	Structure	Value 'n'	LeNCED1 inhibition at 100 µM (%)
A1		--	20
B1*		--	5
B2		--	13
E1		1	16
E2		2	17
E3		3	22
E4		4	9

Table 3.3 – Inhibition data for hydroxamic acid compounds A1, B1, B2 and E1-4 versus LeNCED1. Compound marked with '*' previously published in *Sergeant et al*. 2009.

between the carbonyl of the hydroxamic acid and the terminal phenyl group. A para substituted fluorine on the phenyl group attached to the *N*-alkoxy of the hydroxamic acid offers no change in the level of inhibition. If the para hydroxyl (D1, D2) is substituted with a methoxy group, a slight increase in inhibition is observed. Given the similarity between the inhibition of D1, D2 (para hydroxyl – hydrogen bond donor) and D4 (para methoxy – hydrogen bond acceptor), there is unlikely to be a hydrogen bonding interaction within the active site. Interestingly, dihydroxylation of the carbonyl phenyl group results in a sharp decrease in inhibition to only 4%. This is unexpected, given that hydroxyl groups would donate electron density into the aromatic ring, increasing any π-π stacking interaction. Similarly, with dimethoxylation, the inhibitory effects are perturbed. This suggests there is a steric effect in play at the meta position. However, with compound D7, which has a 3,4-methylenedioxy substitution, the level

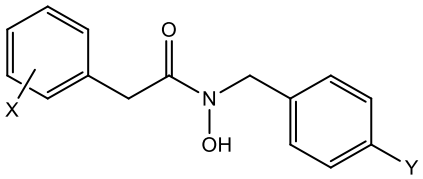
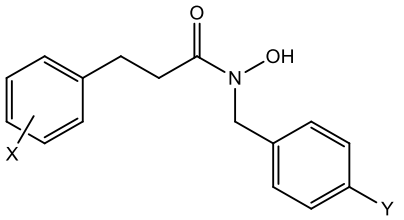
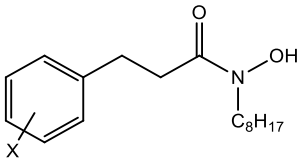
Inhibitor	Structure	Substituent X	Substituent Y	LeNCED1 inhibition at 100 µM (%)
D1*		4-Hydroxy	H	27
D2*		4-Hydroxy	F	29
D3*		3,4-Dihydroxy	F	4
D4*		4-Methoxy	F	33
D5*		3,4-Dimethoxy	H	8
D6*		3,4-Dimethoxy	F	18
D7*		4-Methoxy	F	33
D16		4-*o*-Butyl	H	14
D17		4-Acetyl	H	20
D8*		3,4-Dimethoxy	H	40
D9*		4-Methoxy	H	27
D14		3,4-Dimethoxy	F	N/A
D15[!]		4-Methoxy	F	24
D10*		3,4-Dimethoxy	---	14
D11*		4-Methoxy	---	15

Table 3.4 – Inhibition data for hydroxamic acid compounds D1-D11 and D14-D16 versus LeNCED1. Compounds marked with an asterisk previously published in Sergeant *et al*. 2009. Compound marked with '!' previously published in Van Norman *et al*. 2014.

of inhibition reported is equal to that of D1, D2 and D4. The mediocre levels of inhibition seen for D16 and D17, 14% and 20% respectively, are most likely due to a steric effect of the bulky *O*-butyl and *O*-acetyl groups.

The inhibitors D8, D9, D14 and D15 are structurally similar to D4, D5 and D6, save for the presence of an ethylene bridge between carbonyl and the phenyl group. The presence of this ethylene linker appears to negate the negative effects of meta hydroxylation and methoxylation. D8, which is 3,4-dimethoxy substituted, reported 40% inhibition at 100 μM, the second highest of all the hydroxamic acids tested. The additional chain length may well be positioning the methoxy groups such that they avoid a steric clash, or possibly such that they act as a hydrogen bond acceptor to a backbone amide.

Replacement of the *N*-alkoxy terminal phenyl group (D10, D11) or its removal (D12, D13) has little effect on the levels of inhibition seen. This suggests that the coordination of the hydroxamic acid moiety and possibly π-π stacking of the carbonyl phenyl group is most important for inhibition. The inhibitors D20 and D21 have a carboxylate group on the *N*-alkoxy of the hydroxamic acid and show modest levels of inhibition. There is possibly a hydrogen bonding or electrostatic interaction involving the carboxylate. The inhibition shown by D21 certainly suggests that the carboxylate is not detrimental to inhibition. In the case of D21, the napthyl group could provide additional π-π stacking interactions within the active site of the enzyme.

The inhibitor D30 shows the highest level of inhibition, 42%, at 100 μM. However, two related compounds, D31 and D32 show quite poor inhibition, only 10% at 100 μM. The inhibitors D30, D31 and D32 all have geranyl units on either the *N*-alkoxy or carbonyl of the hydroxamic acid. In compounds D31 and D32 there is an alkene conjugated to the carbonyl group of the hydroxamic acid. In D30, this alkene is one carbon unit further away from the hydroxamic acid and it is on the *N*-alkoxy terminus of the hydroxamic acid, preventing any conjugation. This lack of conjugation allows free rotation around the carbon-nitrogen σ bond.

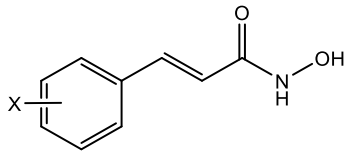
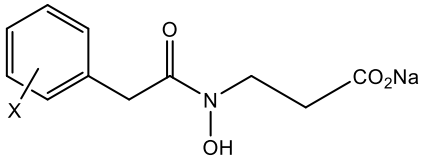
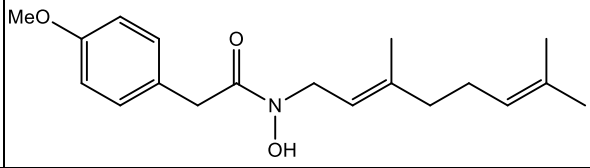
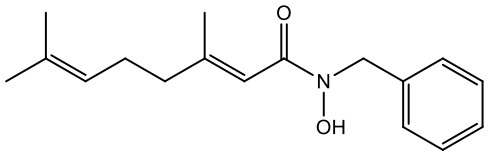
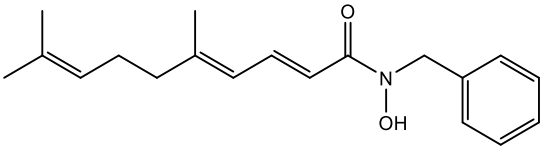
Inhibitor	Structure	Substituent X	LeNCED1 inhibition at 100 µM (%)
D12*		3,4-Dimethoxy	11
D13*		4-Methoxy	13
D20		4-Methoxy	22
D21		Napthyl	39
D30		---	42
D31		---	10
D32		---	10

Table 3.5 – Inhibition data for hydroxamic acid compounds D12-D13, D20-D21 and D30-D31 versus LeNCED1. Compounds marked with an asterisk previously published in Sergeant *et al*. 2009.

This increase in flexibility may well allow the inhibitor to maximise interactions with the active site, aiding inhibition.

Within the F series of compounds, no inhibition greater than 7% was recorded. Compounds in the F series lack a methylene or ethylene linker between the carbonyl moiety of the hydroxamic acid and the phenyl group. One could propose that this lack of a methylene or ethylene bridge prevents the phenyl group from aligning a π-π stacking interaction or other favourable interaction within the active site. Halogen or amino groups substituted onto the meta position of the phenyl group have little effect on inhibition. Given the electron withdrawing effects of halogens, it would be expected that F5 and F7 would show no

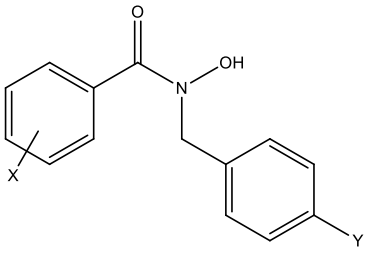
Inhibitor	Structure	Substituent X	Substituent Y	LeNCED1 inhibition at 100 µM (%)
F1*		4-Methoxy	H	0
F2*		4-Methoxy	F	0
F3*		3,4-Dimethoxy	H	2
F4*		3,4-Dimethoxy	F	0
F5		3-Chloro	H	5
F6		3-Amino	H	7
F7		3-Bromo	H	4
F8		3-Chloro-4-Methoxy	H	0

Table 3.6 – Inhibition data for hydroxamic acid compounds F1-F8 versus LeNCED1. Compounds marked with an asterisk previously published in Sergeant *et al*. 2009.

inhibition, not more inhibition. The electron donating effect of the amino group on F6 is possibly helping to improve the inhibition seen by the compound.

Compounds in the 'H' series showed more effective inhibition, generally around 25% at 100 µM. The H series compounds are sterically quite bulky, and contain many π-bonding systems, which may aid their binding to the enzyme. Compound B2, which is structurally similar to the H series compounds, showed similar levels of inhibition at 100 µM. It may also be that the increased chain length aids the flexibility of the inhibitors. Given the NCED is selective for the 9-*cis* isomer of 9'-*cis*-neoxanthin, it may be that the most effective inhibitors in fact mimic this geometrical feature.

The magnitude of the difference between the most effective of the hydroxamic acids against LeNCED1 and the least effective is not great, and does not allow for strong conclusions to be made on the aspect of structure activity relationships (Figure 3.7). One of the clear trend that emerges from this data is that compounds with longer chain lengths appear to show greater inhibition. This could be because these longer chained compounds are maximising hydrophobic binding within the active site of LeNCED1. Looking at the active site of *Z. mays* Vp14, a related NCED enzyme which shares significant sequence homology (66% identity of LeNCED1 with Vp14) (see Figure 1.9, Section 1.4) in Figure 3.8, it can be

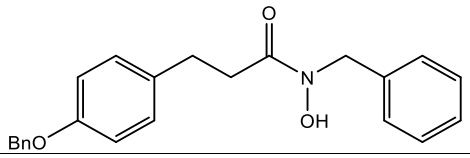
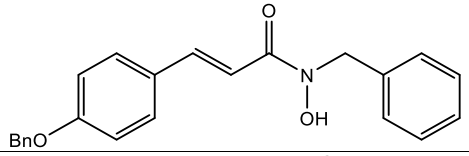
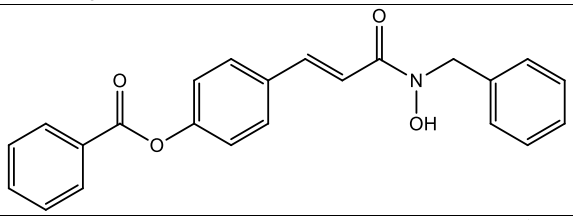
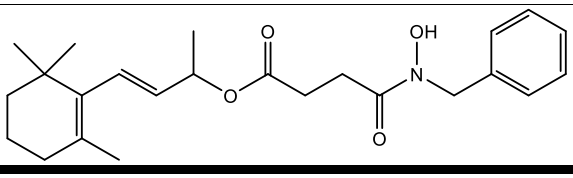
Inhibitor	Structure	LeNCED1 inhibition at 100 µM (%)
H1		12
H2		24
H3		26
H4		24

Table 3.7 – Inhibition data for hydroxamic acid compounds H1-H4 versus LeNCED1. Bn represents a benzyl group (CH_2Ph).

seen that there are a significant number of hydrophobic residues within the active site (Messing *et al*. 2010). However, there are several exceptions to this trend. B2, for example, a long chain hydroxamic acid, shows low inhibition against LeNCED1. Given the structural similarities between D30, D31 and D32, it is interesting that D31 and D32 showed lower

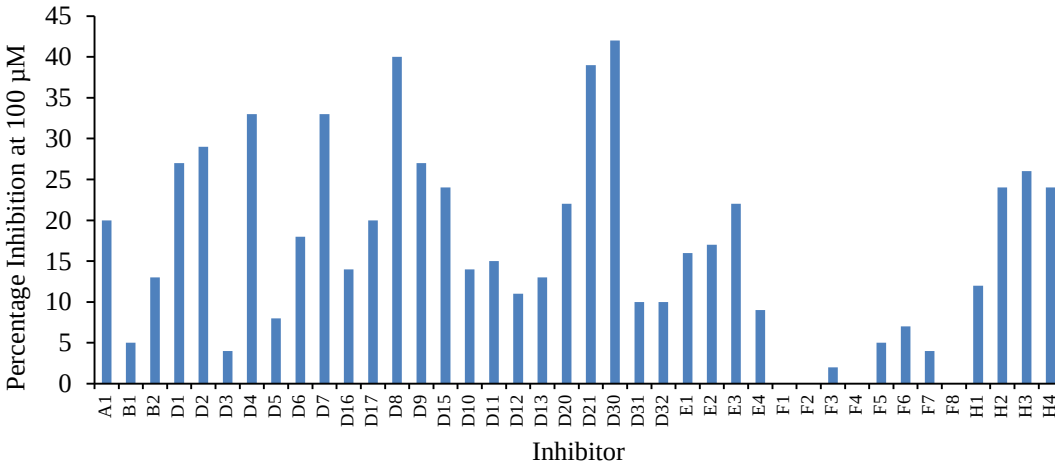


Figure 3.7 – Summary of inhibition shown by hydroxamic acids against cell free extract containing overproduced LeNCED1.

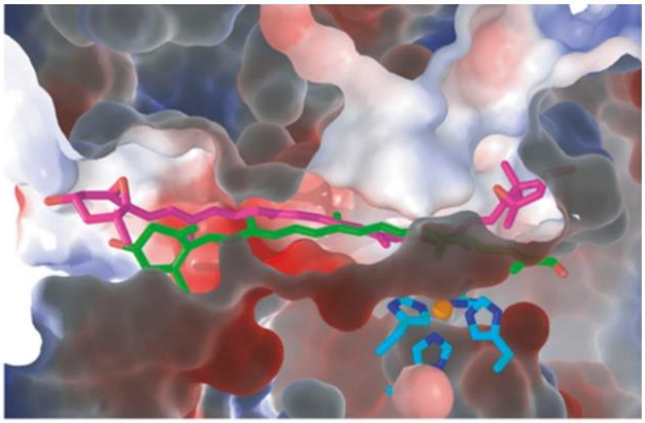


Figure 3.8 – Active site structure of *Z. mays* VP14 with docked 9'-*cis*-violoxanthin (purple) and *β*–apo-8'-carotenal (green) showing the polarities of residues found within the active site. Figure taken from Messing *et al*. (2010).

inhibition than D30. In addition, D10 and D11 both showed weak inhibition despite having a long aliphatic chain.

One particular concern with performing assays with cell free extract is non-specific binding of the inhibitor molecules to other soluble proteins within the cell free extract. However, given that cell free extract was used for *in vitro* assays for LeCCD1a (see Chapter 2.2), and inhibition of greater than 95% was recorded for several inhibitors, it seems unlikely that this is the cause of the low levels of inhibition seen. It seems more likely that the specific geometry and features of the NCED active site are preventing inhibition.

3.4 Expression and Assays of *Z. mays* NCEDs versus Inhibitors D2 and D4

When assayed *in planta* against *A. thaliana*, the hydroxamic acids D2 and D4 (Figure 3.9) exhibit phenotypic effects on seed germination, whereby application of the compounds D2 and D4 to the seed decreases the time taken for the seed to germinate (Sergeant *et al.* unpublished). Seed germination is controlled by the plant signalling hormone abscisic acid,

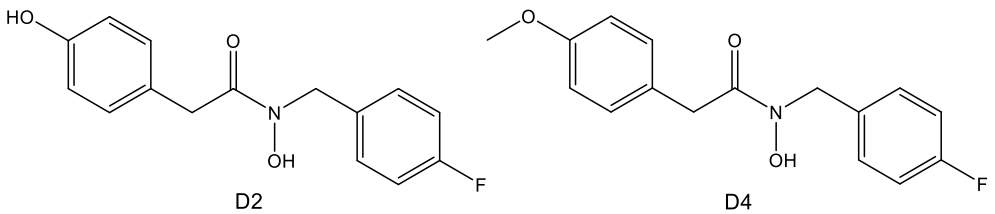


Figure 3.9 – Structures of the hydroxamic acid inhibitors D2 and D4.

the biosynthesis of which is controlled by NCED (Taylor *et al*. 2005). However, within *Arabidopsis*, and other plant species in fact, there are multiple NCED homologs, and it is unknown which NCED enzyme, or enzymes, are the cause of the observed phenotype (Tan *et al*. 2003, Lefebvre *et al*. 2006). *Z. mays* is a commercially important crop which has five different NCED homologues termed ZmNCED1, 2, 3A, 3B and 9 (Chandler 2014). In an attempt to identify which NCED homologue is responsible for the observed phenotype and to aid in the development of a specific NCED inhibitor, it was decided to express the *Z. mays* NCED enzymes and to assay them against the two more effective hydroxamic acid inhibitors, D2 and D4 (provided by Syngenta). In addition, there is not a strong correlation between the levels of inhibition seen *in vitro* against *S. lycopersicum* NCED1 and the *in planta* effect on seed germination.

The *Z. mays* NCEDs were cloned by Jake Chandler (Cranfield University) from *Z. mays* cDNA into the pGEX-4T-1 vector (Chandler 2014). The ZmNCEDs were then used to transform *E. coli* BL21 and expressed (Figure 3.10). Sequences of the five ZmNCEDs are shown in appendix 11.2. Three ZmNCEDs, ZmNCED1, 3A and 3B expressed moderately well, considering the codon bias difference between *Z. mays* and *E. coli* (Tuller *et al*. 2010, Burgess-Brown *et al*. 2008). A fourth, ZmNCED2, expressed at a low level, however, there was enough protein present to allow for an assay. For the fifth NCED, ZmNCED9, no protein band was visible at all. ZmNCED9 was then expressed in pRosetta BL21 *E. coli* in an attempt to enhance the expression, since the pRosetta plasmid codes for rare codons which are not usually required for gene expression in *E. coli* (Kudla *et al*. 2009, Fu *et al*. 2007, Kane 1995).

Using cell free extracts containing overexpressed ZmNCED, the inhibitors D2 and D4 were assayed at 100 µM concentrations. Assays were conducted as described for *S. lycopersicum* NCED1 and the products were examined by HPLC as described for *S. lycopersicum* NCED1. Inhibition data for ZmNCED1, 2, 3A and 3B is shown in Figure 3.11. Despite also being expressed in pRosetta BL21 *E. coli* strains no turnover of the 9'-*cis*-neoxanthin substrate was observed with the cell free extract. Against NCED1, 2, 3A

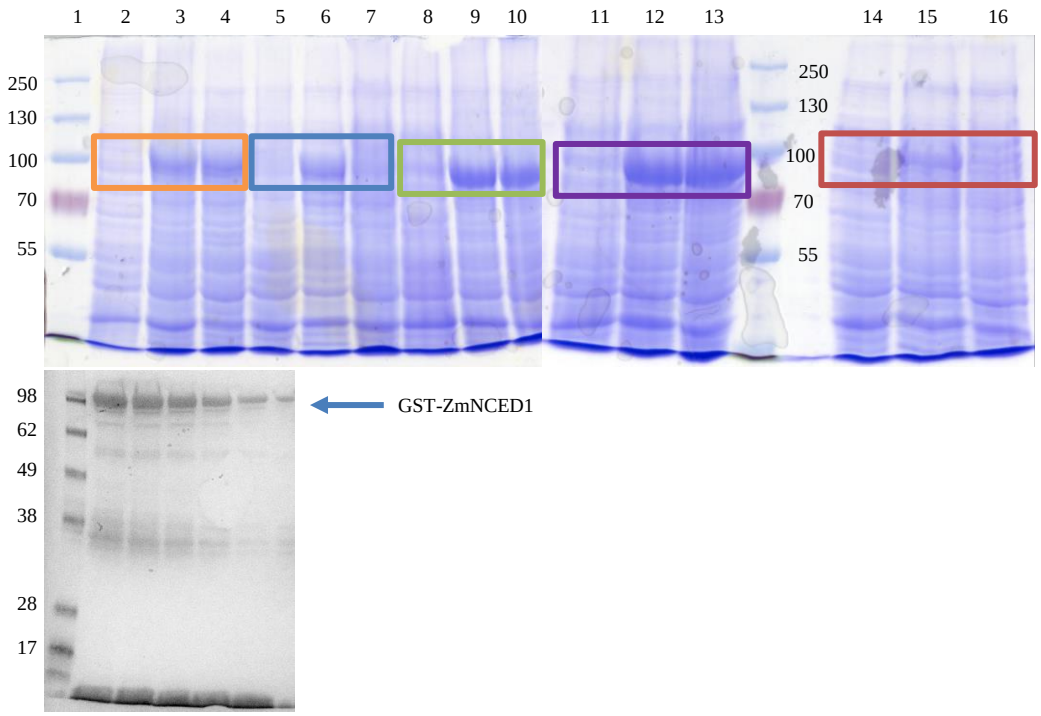


Figure 3.10 – Top: 8% SDS-PAGE analysis of the production of *Z. mays* enzymes NCED1 (orange box), 2 (blue box), 3A (green box), 3B (purple box) and 9 (red box) expressed in *E. coli* BL21. Thick band between 70 and 100 kDa marker corresponds to the overproduced NCED. 1 – Ladder; 2 - ZmNCED1 non-induced control (NIC); 3 – ZmNCED1 total protein (TP); 4 – ZmNCED1 total soluble protein (TSP); 5 – ZmNCED2 NIC; 6 – ZmNCED2 TP; 7 – ZmNCED2 TSP; 8 – ZmNCED3A NIC; 9 – ZmNCED3A TP; 10 – ZmNCED3A TSP; 11 – ZmNCED3B NIC; 12 - ZmNCED3B TP; 13 – ZmNCED3B TSP; 14 – ZmNCED9 NIC; 15 – ZmNCED9 TP; 16 – ZmNCED9 TSP. Molecular weights of GST-NCEDs (calculated from sequence); 1: 89.2 kDa; 2: 86.9 kDa; 3A: 83.3 kDa; 3B: 83.0 kDa; 9: 89.0 kDa. Molecular weights of ladder are in kDa. Mass spectrometry would be required to definitely confirm the production of each protein. **Bottom:** Western blot analysis of ZmNCED1 elution fractions following GST purification visualised with anti-GST-HRP conjugate using BioRad ChemiDoc imaging instrument.

and 3B both D2 and D4 showed 20-40% inhibition, apart from D2 versus ZmNCED2, which showed approximately 65% inhibition. However, if one particular isozyme was responsible for the observed phenotype at 1 mM concentration *in planta*, then one might expect greater than 90% inhibition for that enzyme at 100 µM of inhibitor *in vitro*, but that effect was not observed. The expression levels of the different NCEDs vary between different plant tissues, and as such, given the phenotypes observed *in planta*, it was expected there would be different levels of inhibition observed *in vitro* (Tan *et al*. 2003, Lefebvre *et al*. 2006). This was not apparent and further work may be required to solubilise ZmNCED9 and to observe what levels of inhibition this enzyme shows against compounds D2 and D4.

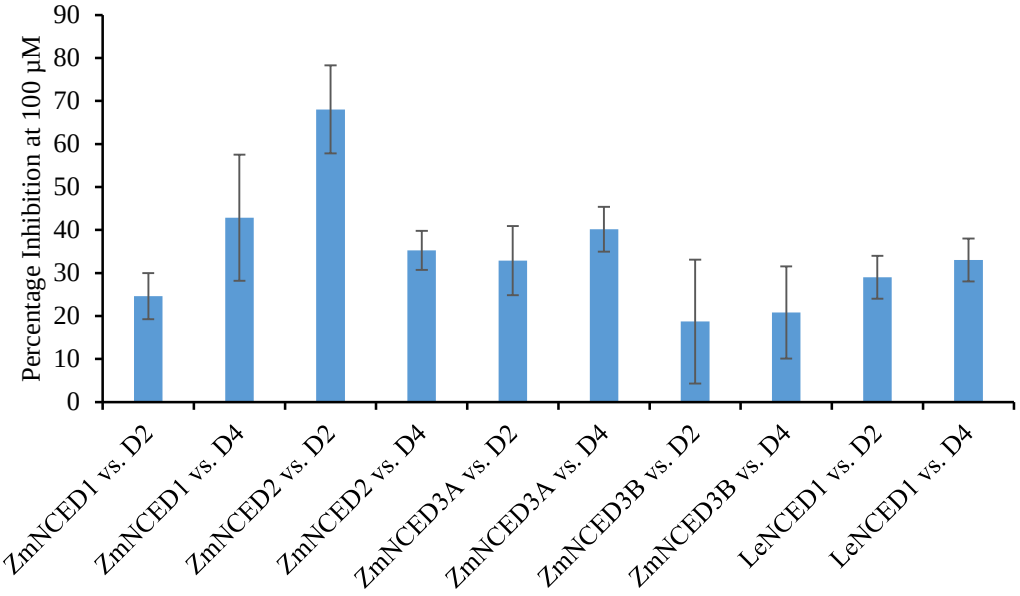


Figure 3.11 – Percentage inhibitions shown by the hydroxamic acid inhibitors D2 and D4 against *Z. mays* NCED1, NCED2, NCED3A, NCED3B, NCED9 and *S. lycopersicum* NCED1. Error bars represent standard deviation of inhibition between replicates.

3.5 Time-Dependent Inhibition of *Z. mays* NCED1

When ZmNCED1 was expressed and purified for the purposes of the development of a coupled NCED-ABA2 assay (see Chapter 3.6), it was discovered that pre-incubation of inhibitors with cell free extract containing overexpressed ZmNCED1, prior to the addition of the substrate, resulted in different levels of inhibition than was observed without pre-incubation. In samples that were pre-incubated, almost 100% inhibition of the enzyme was observed (Figure 3.12). Conversely, in samples without pre-incubation, inhibition levels of less than 40% were seen (Figure 3.12). On the basis of these observations, *Z. mays* NCED1

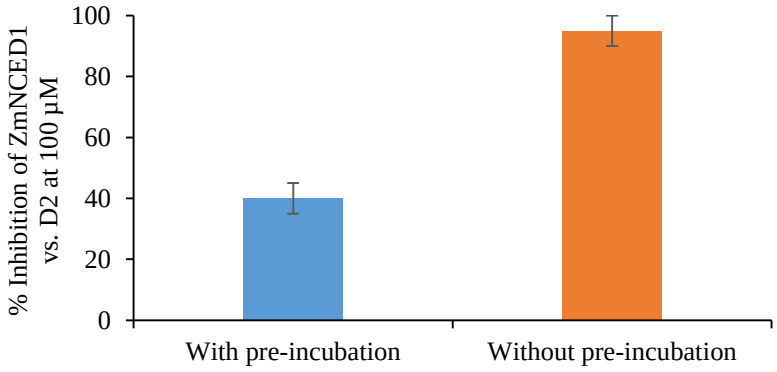


Figure 3.12 – Comparison of percentage inhibition observed at 100 µM D2 versus cell free extract containing over produced ZmNCED1 with and without pre-incubation of D2.

was purified via the same method as reported for *S. lycopersicum* CCD1. Reactivated *Z. mays* NCED1 was shown to effectively turn over 9'-*cis*-neoxanthin to xanthoxin and the C_{25} aldehyde via HPLC. However, when D2 was pre-incubated with the enzyme for different lengths of time, either 30, 20 or 10 minutes, no turnover was observed (Figure 3.13). Partial inhibition was seen when D2 was added simultaneously with the substrate. This time dependent inhibition effect has been seen with hydroxamic acid inhibitors of *p*-hydroxy-phenylpyruvate dioxygenase (HPPD) (Sergeant *et al*. 2013). In this case the 'F-series' inhibitors were shown to inhibit HPPD to a greater extent on pre-incubation of the inhibitor with the enzyme. This effect was speculated to be due to oxidation of the Fe (II)-HA complex to Fe (III)-HA (where HA represents the hydroxamic acid). Fe (III) would bind the hydroxamic acid more strongly, thus resulting in the increased levels of inhibition observed (Sergeant *et al*. 2013).

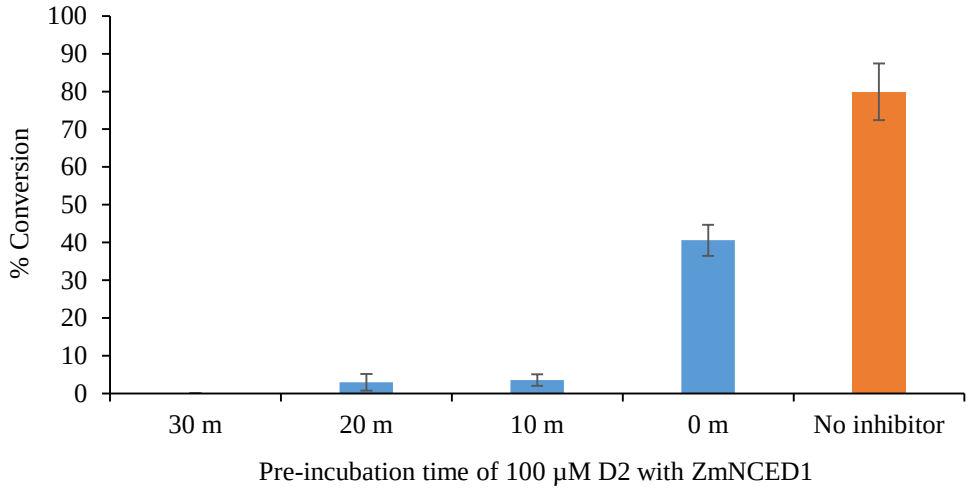


Figure 3.13 – Comparison of percentage conversion of 9'-*cis*-neoxanthin to C_{25} aldehyde by ZmNCED1 following pre-incubation of ZmNCED1 with compound D2 for various lengths of time.

Given the effect of pre-incubation of the hydroxamic acids D2 and D4 with ZmNCED1 it appears that the biochemical basis of the seed germination phenotype observed in *Z. mays* is due to inhibition of one or more NCED isozymes. Interestingly, however, when the inhibitors abamine and abamineSG (provided by Syngenta) are assayed against ZmNCED1 using the same time dependent incubation method, only 20% inhibition is observed at 100 µM (Figure 3.14). This is interesting, given that abamine and abamineSG

have been shown to inhibit stomatal closure, ABA induced gene expression of *RD29B* (a gene with an ABA response element which encodes the LEA protein in *A. thaliana*) and osmotic induced ABA accumulation *in planta* in *A. thaliana* (Han *et al*. 2004, Kitahata *et al*. 2006). AbamineSG has additionally been shown to inhibit seedling growth in *A. thaliana* (Kitahata *et al*. 2006). As with the hydroxamic acids, complete inhibition would be expected at 100 µM *in vitro* if these phenotypes were the result of NCED inhibition *in planta*. This would suggest that there is perhaps an additional target for these compounds. Some hydroxamic acids (D12 and D13) have been shown to inhibit other iron-dependent dioxygenases such as the *E. coli* catechol dioxygenase MhpB and HPPD (Sergeant *et al*. 2013, Sainsbury 2014). As such, the hydroxamic acids could be inhibiting another non-heme oxygenase enzyme resulting in the observed phenotype.

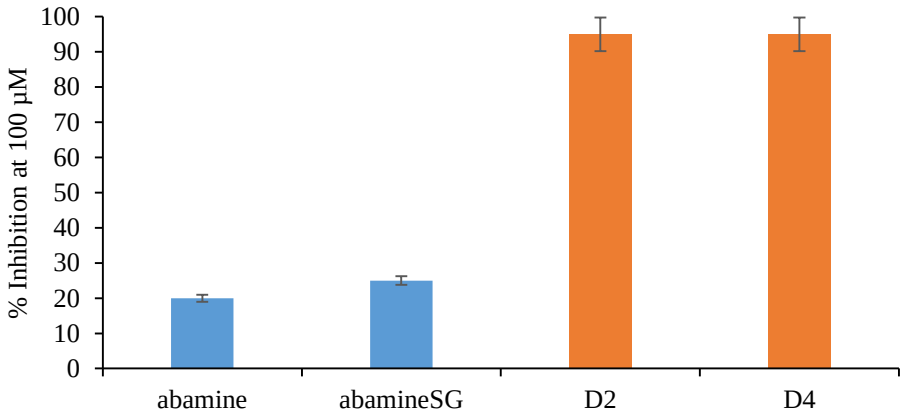


Figure 3.14 – Comparison of percentage inhibition observed for abamine, abamineSG, D2 and D4 at 100 µM against ZmNCED1 following 10 minute pre-incubation of ZmNCED1 with inhibitor prior to the addition of 9'-*cis*-neoxanthin.

3.6 Development of a Coupled NCED-ABA2 Assay

The rate limiting step in the biosynthesis of ABA is the 11',12' cleavage of 9'-*cis*-neoxanthin catalysed by NCED (Taylor *et al*. 2005). As explained in Section 3.2, the NCED enzyme and its 9'-*cis*-neoxanthin substrate present many problems when being assayed. HPLC assays, required for the study of NCED, require significantly more time for sample preparation, data collection and analysis. As a result, this makes the screening of large collections of compounds rather difficult and time consuming. To this end, it is necessary to

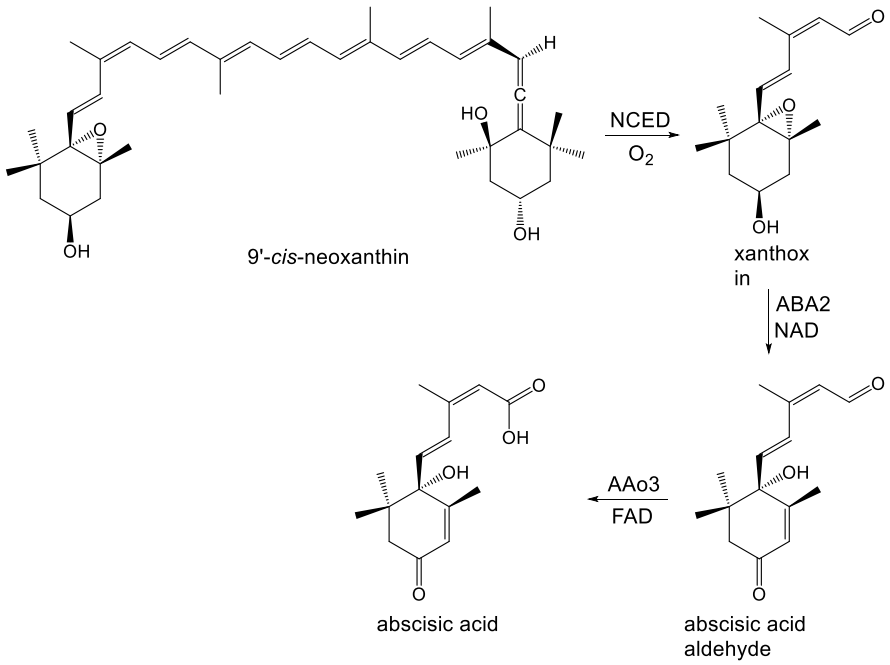


Figure 3.15 – Biosynthetic route from 9'-*cis*-neoxanthin to the signalling hormone abscisic acid.

develop a high throughput assay for the screening of compounds for inhibition of NCED in order to identify an effective and selective NCED inhibitor.

Xanthoxin is, along with the C_{25} aldehyde, produced upon NCED cleavage of 9'-*cis*-neoxanthin (Figure 3.15). Xanthoxin is a conjugated aldehyde which is oxidised to abscisic aldehyde by ABA2, also known as xanthoxin dehydrogenase, the next enzyme on the ABA biosynthesis pathway (Figure 3.15) (González-Guzmán *et al*. 2002, Cheng *et al*. 2002). The enzyme abscisic aldehyde oxidase (AAo3) then catalyses the final step in the biosynthesis of abscisic acid (Seo *et al*. 2000). ABA2 is an alcohol dehydrogenase and belongs to a family of enzymes known as the short chain dehydrogenases (González-Guzmán *et al*. 2002, Cheng *et al*. 2002). The oxidation of xanthoxin by ABA2 is a nicotinamide adenine dinucleotide (NAD) dependent process, and as such the reaction can be monitored by the change in absorbance or fluorescence at 340 nm or 440 nm respectively as reduced NAD (NADH) is produced. If ABA2 were coupled to NCED, then as NCED produces xanthoxin, it should be oxidised by the ABA2, resulting in the production of NADH, which can be monitored continuously. The reported K_M for ABA2, of 19 µM, is lower than reported K_M values for NCEDs (49 µM, 27

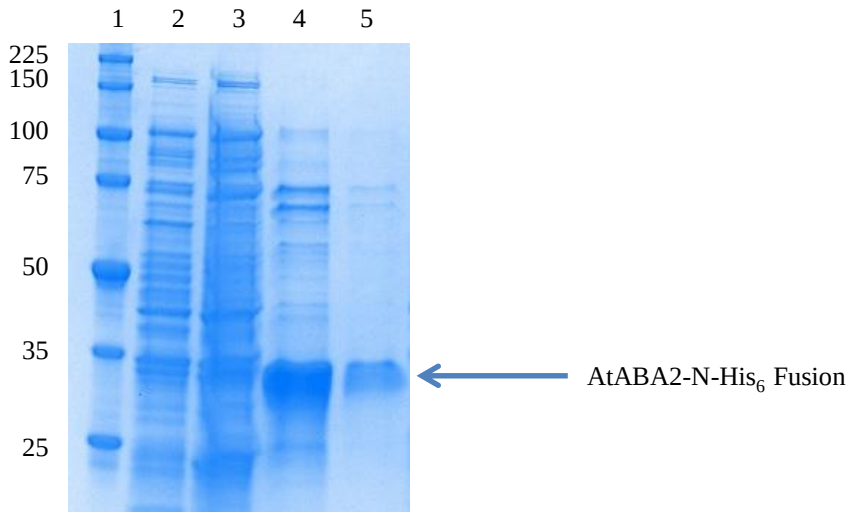


Figure 3.16 – SDS-PAGE gel of N-His_6-AtABA2 purification. Lane 1 – Ladder; 2 – Non-induced *E. coli* cell pellet; 3 - Induced pET-200-AtABA2 *E. coli* cell pellet; 4 – Elution fraction 1; 5 - Elution fraction 3. Molecular weights of the ladder are in kDa. Molecular weight of N-His_6-AtABA2 (calculated from sequence): 34.4 kDa. Mass spectrometry would be required to definitively confirm expression of AtABA2.

µM and 24 µM for *V. unguiculata* NCED, *Z. mays* NCED1 and *A. thaliana* NCED3 respectively (Han *et al*. 2004, Schwartz *et al*. 2003, Boyd *et al*. 2009).

To this end, the *A. thaliana* ABA2 gene sequence was codon optimised for expression in *E. coli*, and the codon optimised gene was purchased from Genscript (see Appendix 10.3 for the optimised gene sequence. Accession number of wild type ABA2 gene: BT003412). The ABA2 gene was cloned from the pUC57 vector into a pET-200 expression vector with an N-His_6 fusion tag and the AtABA2 gene fusion product was expressed in BL21 *E. coli* and purified by nickel affinity chromatography (Figure 3.16). Work by González-Guzmán *et al*. has shown that ABA2 will also oxidise a range of other alcohols, including 3, 3, 5-trimethylcyclohexanol (González-Guzmán *et al*. 2002). As such 3, 3, 5-trimethylcyclohexanol was used as a control to test the activity of AtABA2 *in vitro*. Upon the addition of 100 µM 3, 3, 5-trimethylcycohexanol in the presence of 100 µM NAD^+ and 36 µg mL^{-1} AtABA2 there was an increase in absorbance at 340 nm, indicating AtABA2 activity (Figure 3.17A). Additionally, NADH production was confirmed by fluorescence under the same conditions (Figure 3.17B). A time dependent increase in emission at 440 nm

can be observed for the AtABA2 catalysed reaction. Assays in the presence of boiled, denatured enzyme and in the absence of NAD showed no turnover (data not shown).

Initial experiments towards the coupled assay involved using cell free extract containing overproduced *S. lycopersicum* NCED1. It was found that upon the addition of the cell free extract containing overproduced LeNCED1 to 100 µM NAD^+, 36 µg mL^{-1} AtABA2 and 40 µM 9'-*cis*-neoxanthin there was an increase in absorbance at 340 nm, consistent with the apparent oxidation of xanthoxin by ABA2 (Figure 3.18A). However, subsequent experiments showed inconsistent results and controls with cell-free extract without

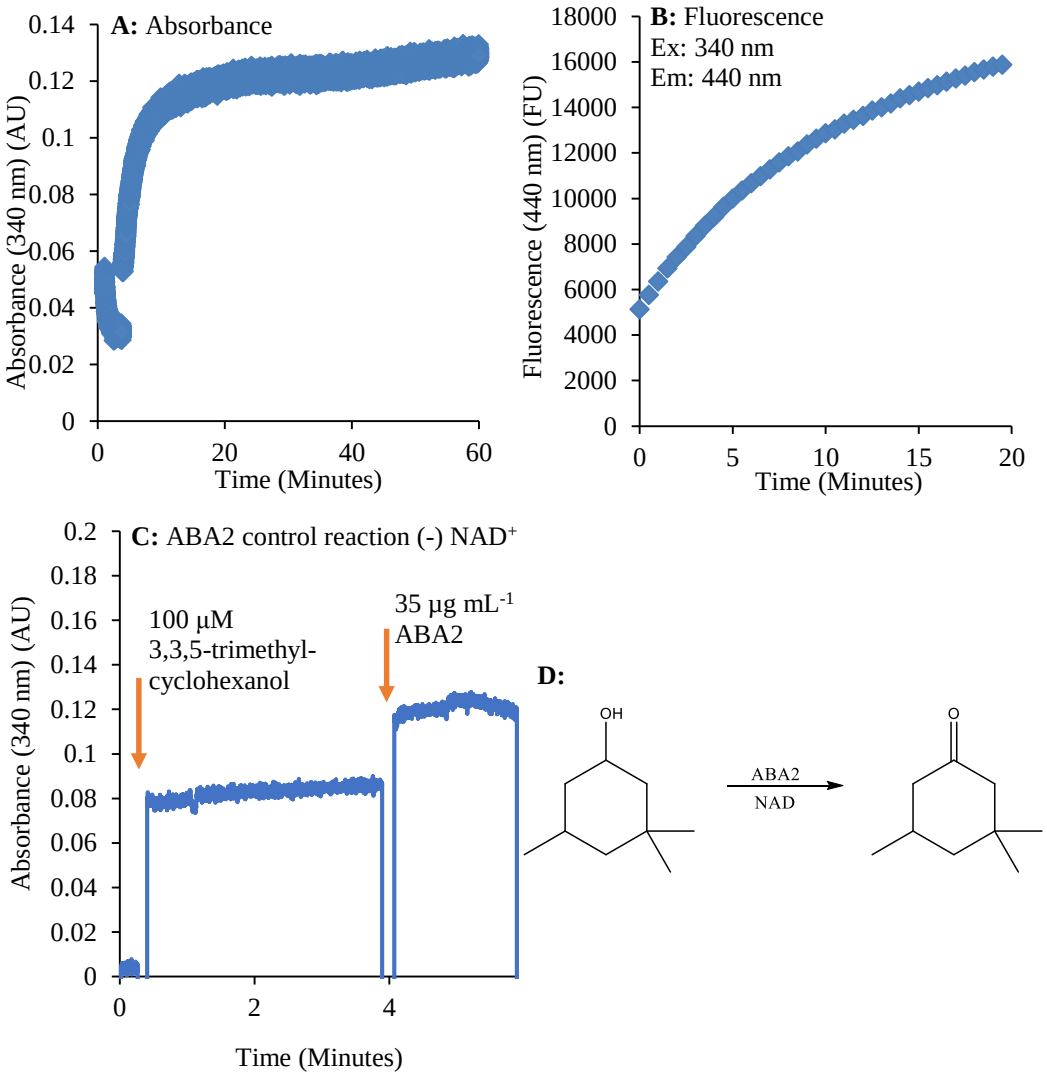


Figure 3.17 – Changes in absorbance (**A**) and fluorescence (**B**) seen for the AtABA2 catalysed oxidation of 3,3,5-trimethylcyclohexanol (**D**). (**C**) ABA2 control reaction in the absence of NAD indicates reaction is dependent upon NAD^+ and an NAD^+ dependent enzyme. Ex – Excitation wavelength; Em: Emission wavelength.

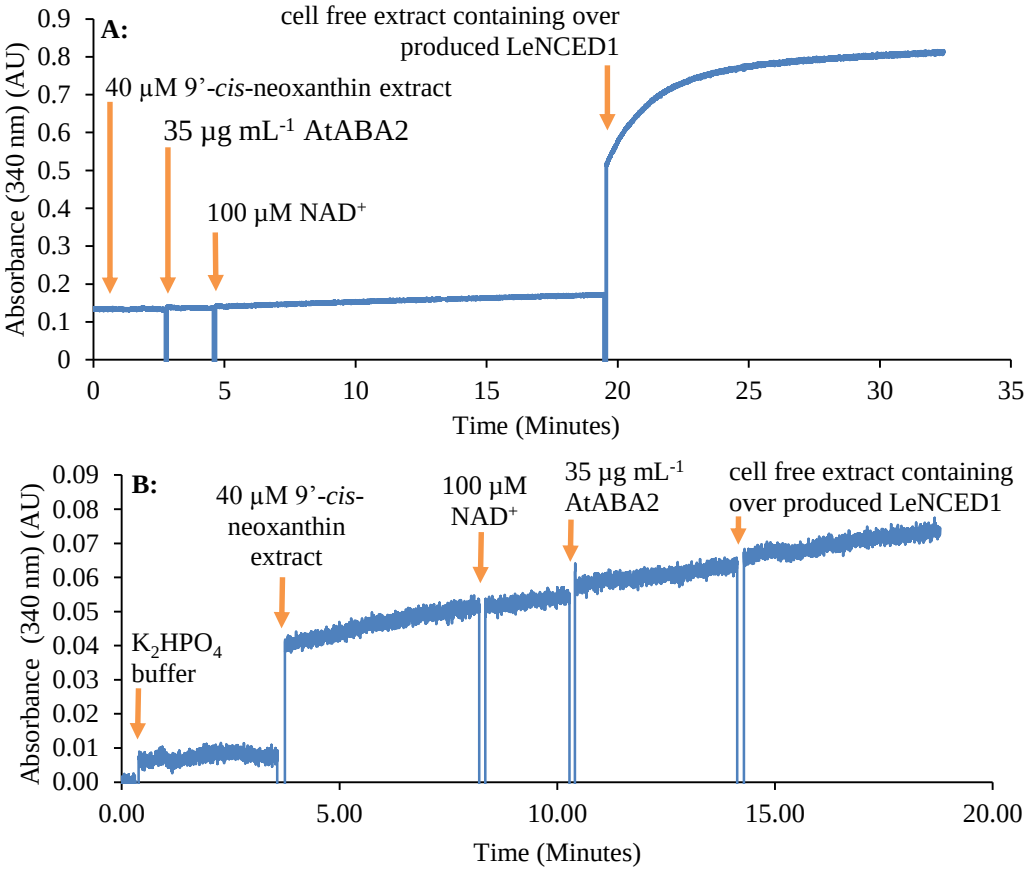


Figure 3.18 – Changes in absorbance at 340 nm during coupled assay development. Top graph, **A**, shows initial attempts towards the development of the coupled assay with an increase in absorbance at 340 nm seen on addition of cell free extract containing over-produced LeNCED1. Bottom graph, **B**, shows the increase in absorbance observed upon the addition of 9'-*cis*-neoxanthin extract to the assay.

overproduced LeNCED1 also caused an increase in absorbance at 340 nm. It was noticed that upon the addition of the 9'-*cis*-neoxanthin extract to the assay mixture there was an increase in absorbance which continued after the addition of the cell free extract (Figure 3.18B). These results suggested that the increase in absorbance was not due to xanthoxin oxidation by ABA2. The upwards drift begins before the addition of ABA2, suggesting it was not due to NAD(H). The possibility that the ABA2 was oxidising the ethanol that is required to solubilise the 9'-*cis*-neoxanthin was considered. However, this was deemed unlikely since experiments by Gonzalez-Guzman *et al* show that AtABA2 does not oxidise ethanol (González-Guzmán *et al*. 2002). The data suggests that something in the 9'-*cis*-neoxanthin extract, the AtABA2 or

the cell lysate was causing a reaction. As a result of the upward drift in absorbance, it was decided to attempt to prepare crude purification samples of the NCED enzyme.

A crude purification of *Z. mays* NCED1 was performed in order to eliminate factors in the cell free extract which could interfere with the assay. ZmNCED1 was chosen instead of LeNCED1 since its purification has been reported in the literature, although its activity was not reported after purification (Messing *et al*. 2010). Purification of both native (with the iron cofactor within the enzyme) and apo-ZmNCED1 (where the iron has been extract via the addition of 1 mM 1, 10–phenanthroline) by glutathione affinity chromatography was performed (via the procedure detailed in Chapter 2.2). The purified proteins were assayed via HPLC under the same conditions as Chapter 3.2. This showed the presence of the C_{25} product, indicating that the ZmNCED1 was active after purification (as seen in Figure 3.5). Due to the crude nature of the purification and the presence of non-specifically bound proteins in the eluted protein fractions, as can be seen by SDS-PAGE gel electrophoresis (for example Figure 3.4), the concentration of ZmNCED1 is unknown, although approximately 50 µg of protein was added to assays. By comparison with HPLC assays involving cell free extract containing over produced ZmNCED1, the activity of the enzyme is diminished after purification. Upon addition of purified native enzyme (5 µL of crude purification elution to a total volume of 200 µL), there was a decrease in absorbance, indicating that the assay was not functional (Figure 3.19A). It is also not clear what would cause this drop in absorbance. When apo-ZmNCED1 was reactivated by the addition of 1.25 mM iron (II) sulphate and 5 mM 2-mercaptoethanol (BME), an increase in absorbance was observed at 340 nm on the addition of the reactivated ZmNCED1 (5 µL of crude purification elution to a total volume of 200 µL) to the assay (Figure 3.19B). Control experiments with boiled enzyme and in the presence of 100 mM EDTA both failed to produce an increase in absorbance at 340 nm (Figure 3.19C and Figure 3.19D). However, control experiments in the absence of NAD resulted in an increase in absorbance at 340 nm (Figure 3.19E). As such, it was not possible to discern whether the ABA2 enzyme was active or not and the cause of this observation remains unknown. One

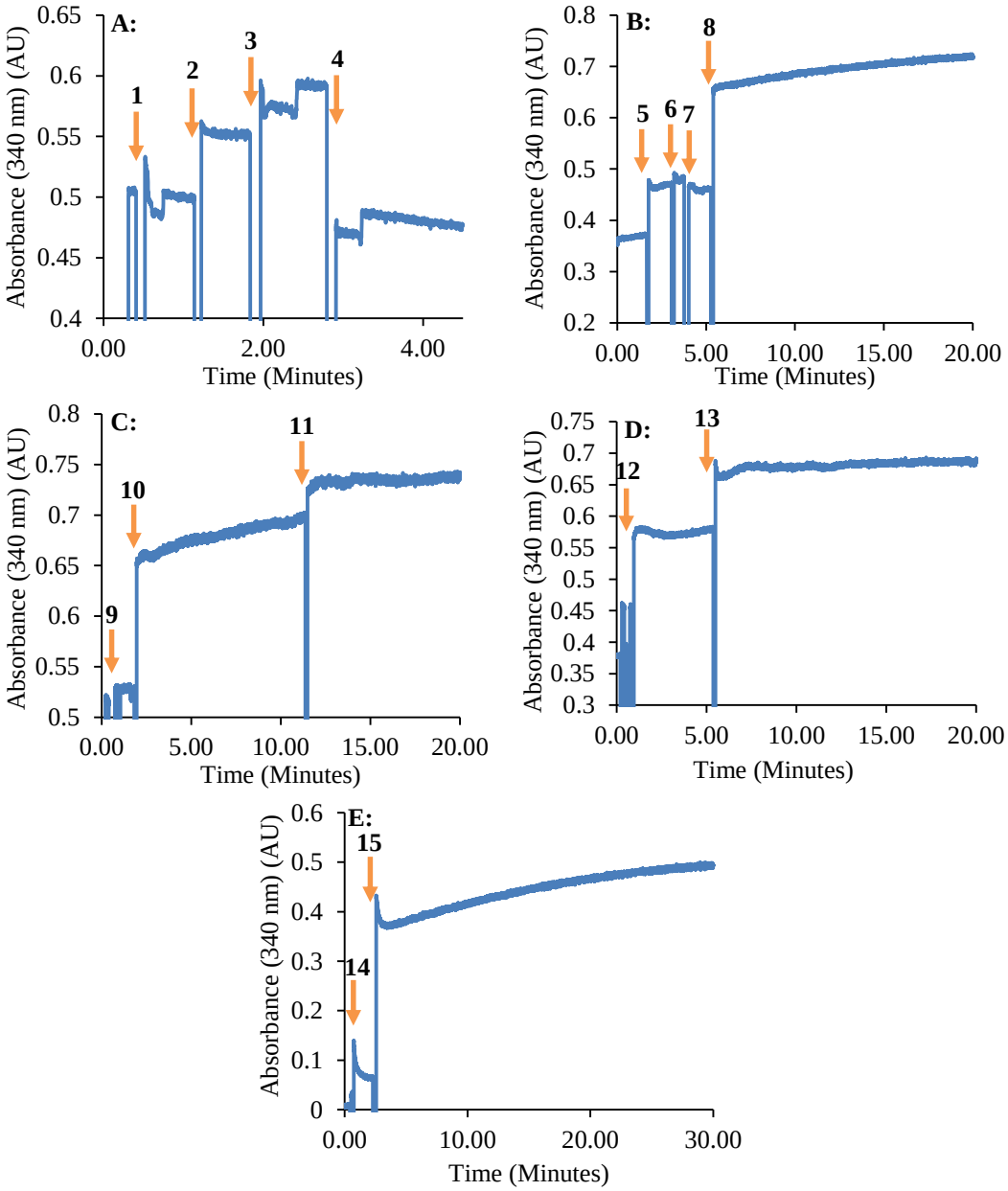


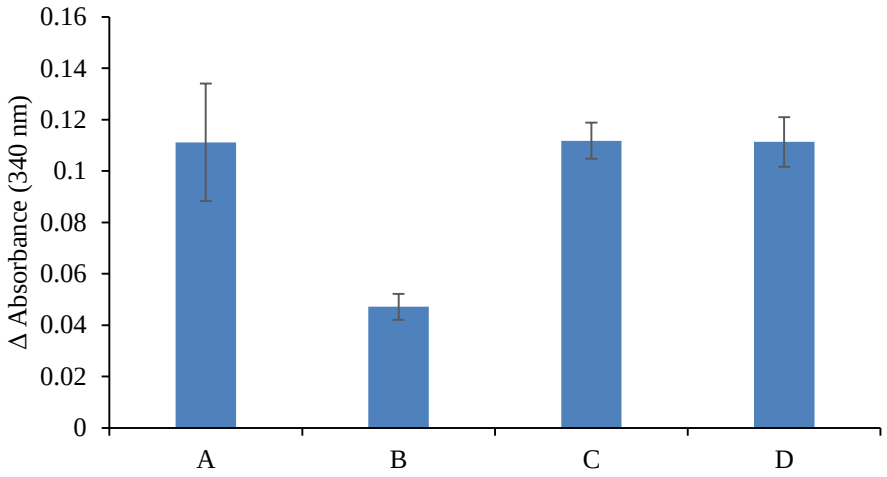
Figure 3.19 – UV chromatograms at 340 nm of NCED-ABA2 coupled assay. Top left, **A**: coupled assay with purified native ZmNCED1. **1** – addition of 100 µM NAD^+; **2** – addition of 35 µg mL^{-1} AtABA2; **3** – addition of 5 µL 9'-*cis*-neoxanthin extract; **4** – addition of 5 µL purified native ZmNCED1. Top right, **B**: coupled assay with reactivated purified apo-ZmNCED1. **5** - addition of 100 µM NAD^+; **6** – addition of 35 µg mL^{-1} AtABA2; **7** – addition of 5 µL 9'-*cis*-neoxanthin extract; **8** – addition of 5 µL reactivated purified apo-ZmNCED1. Middle left, **C**: coupled assay in the presence of EDTA. **9** – addition of 100 µM NAD^+, 35 µg mL^{-1} AtABA2 and 5 µL 9'-*cis*-neoxanthin extract; **10** - addition of 5 µL reactivated purified apo-ZmNCED1; **11** – addition of 5 mM EDTA. Middle right, **D**: coupled assay with boiled reactivated purified apo-ZmNCED1. **12** - addition of 100 µM NAD^+, 35 µg mL^{-1} AtABA2 and 5 µL 9'-*cis*-neoxanthin extract; **13** – addition of 5µL boiled reactivated purified apo-ZmNCED1. Bottom, **E**: coupled assay in the absence of NAD^+. **14** – addition of 35 µg mL^{-1} AtABA2 and 5 µL 9'-*cis*-neoxanthin extract; **15** – addition of 5 µL reactivated purified aop-ZmNCED1.

possible explanation for the results observed is that ABA2 is oxidising the 9'-*cis*-neoxanthin directly before xanthoxin is produced from 9'-*cis*-neoxanthin. This is possible, given that structurally there is no difference between the portion of the ABA2 substrate that is oxidised between xanthoxin and 9'-*cis*-neoxanthin, but Figure 3.19E shows that the increase in absorbance is not due to NAD^+ dependent enzyme activity.

Since it was speculated that interference in the assay was coming from the crude 9'-*cis*-neoxanthin preparation, authentic, pure 9'-*cis*-neoxanthin from CaroteNature was obtained and used. Using a continuous UV assay at 340 nm, 25 µM 9'-*cis*-neoxanthin was added to a sample containing 5 µg mL^{-1} purified ABA2 and 100 µM NAD^+. Upon addition of the NCED there was no significant increase in the absorbance but a small increase of 0.002 AU was recorded over 15 minutes. However, considering the molar extinction coefficient of NADH is 6220 M^{-1} cm^{-1}, and that the 9'-*cis*-neoxanthin is used at a final concentration of 25 µM in the assays, then an increase in absorbance of 0.15 AU at 340 nm would be expected for the coupled assay, given there is a 1:1:1 ratio between the stoichiometry of 9'-*cis*-neoxanthin, xanthoxin and NAD(H) required for the reaction. Hence the observed absorbance change would correspond to only a 1.3% turnover of 9'-*cis*-neoxanthin. Considering the abscisic acid biosynthesis pathway, it is reasonable to assume that ABA2 would not oxidise 9'-*cis*-neoxanthin, as if ABA2 were able to oxidise 9'-*cis*-neoxanthin, then the ABA biosynthesis pathway would have evolved the ABA2 to act before the NCED cleavage step. In addition given the hydrophobic nature of 9'-*cis*-neoxanthin, which is membrane associated, ABA2 would also have had to evolve to act on a hydrophobic substrate at the interface of membranes. Given the highly soluble cytosolic nature of ABA2 this seems unlikely. On the other hand, structurally, oxidised 9'-*cis*-neoxanthin would not be dissimilar to 9-*cis*-neoxanthin, as discussed above.

Using pure 9'-*cis*-neoxanthin and purified ZmNCED, the coupled assay was repeated and the change in absorbance at 340 nm was recorded. In the presence of 100 µM NAD^+, 5 µg mL^{-1} ABA2, 25 µM 9'-*cis*-neoxanthin and 54 µg mL^{-1} of protein containing NCED there

was an increase in absorbance at 340 nm, However, control experiments also showed increases in absorbance by the same magnitude (Figure 3.20). Only one control, NAD, 9'-*cis*-neoxanthin and NCED showed no increase in absorbance. Repeating the assay with cell free extract containing over produced NCED in order to compensate for a possible low concentration of NCED had no effect.



	A	B	C	D
NAD^+	+	+	+	+
AtABA2	+	+	+	
9'-*cis*-neoxanthin	+		+	+
ZmNCED1	+	+		+

Figure 3.20 – Changes in absorbance associated with reduction of NAD^+ shown by the full coupled assay (**A**) and controls (**B**-**D**) at 340 nm. Error bars represent standard deviation between replicates.

As already mentioned, the fluorescence signal of NADH can be monitored at 440 nm which is significantly more sensitive than that of the absorbance signal. However, it was not possible to use NADH fluorescence to monitor the reaction, since the emission of NADH, 440 nm, overlaps with the absorption of the 9'-*cis*-neoxanthin, which is also 440 nm. As such, any signal produced from the reduction of NAD would be quenched by 9'-*cis*-neoxanthin. The fluorescence of carotenoids has been reported (Frank & Christensen 2008), but the quantum yield of the fluorescence signal is low, and as such it is unlikely that any signal could be detected and discerned from a no doubt overlapping signal from the C_{25} product. The fluorescence of NADH can be used for control experiments in the absence of 9'-*cis*-neoxanthin. Accordingly, a series of controls in the absence of 9'-*cis*-neoxanthin were run in

order to determine if either ABA2 or NCED were reacting in some way to reduce NAD. As can be seen in Figure 3.21, none of the control assays showed any significant change in fluorescence at an emission of 440 nm. As such, these components, either on their own or reacting together, are not producing reduced NADH. The result of this is that background absorbance at 340 nm must be either due to 9'-*cis*-neoxanthin or another unknown factor.

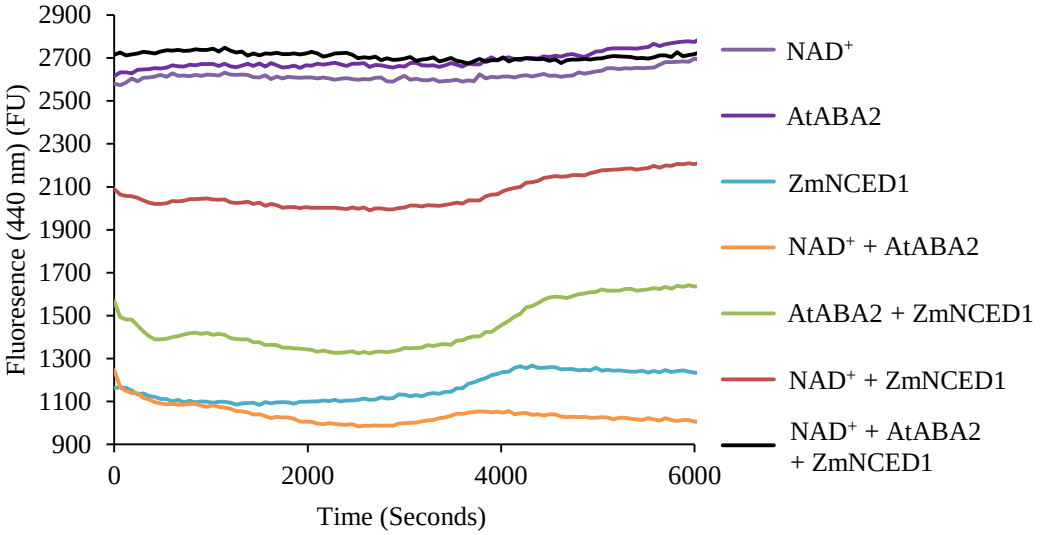


Figure 3.21 – Changes in fluorescence observed at 440 nm for control reactions in the AtABA2-ZmNCED1 coupled assay. Concentrations of 100 µM NAD^+, 5 µg mL^{-1} ABA2, 25 µM 9'-*cis*-neoxanthin and 54 µg mL^{-1} NCED were used for the assays.

In general, assays followed the same pattern shown in Figure 3.20, whereby the absorbance change shown by the controls is equal to that of the full coupled assay. In an attempt to overcome this background absorbance several changes were made to the assay, such as increasing the concentration of NCED used, pre-incubation of NCED with 9'-*cis*-neoxanthin and the removal of factors which were found to cause high background, such as catalase. It is interesting to note that the control without 9'-*cis*-neoxanthin had a lower absorbance change than the controls with 9'-*cis*-neoxanthin (Figure 3.20). This suggests that 9'-*cis*-neoxanthin is causing the absorbance change seen through a non-enzyme catalysed process.

The production of NADH can, however, be monitored using an NAD(P)H-Glo kit (Promega). With this system, the NADH is re-oxidised by a third enzyme, a reductase. This

reductase reduces a substrate to produce luciferin, which is then used by a luciferase enzyme to produce a light signal (Figure 3.22). This signal can then be quantified as the amount of light produced is proportional to the amount of NADH present in the sample. In addition, the NAD(P)H system is sensitive to a concentration of 0.15 μM of NADH (Promega Ltd 2014). A concentration of 25 μM of 9'-*cis*-neoxanthin was used in the assays and even if neither ABA2 and NCED worked at optimal rates, then a signal should still be observed with the NAD(P)H kit. When analysed with the NAD(P)H Glo kit, it was seen that both the coupled assay and the controls all showed high luminescence, suggesting that NADH is being produced in the control reactions (Figure 3.23). However, this contradicts the observations of the control reactions performed using fluorescence, which suggested that the noise at 340 nm was not due to NADH.

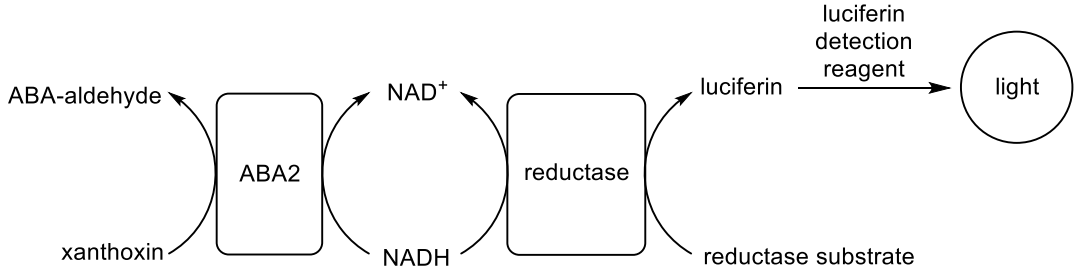
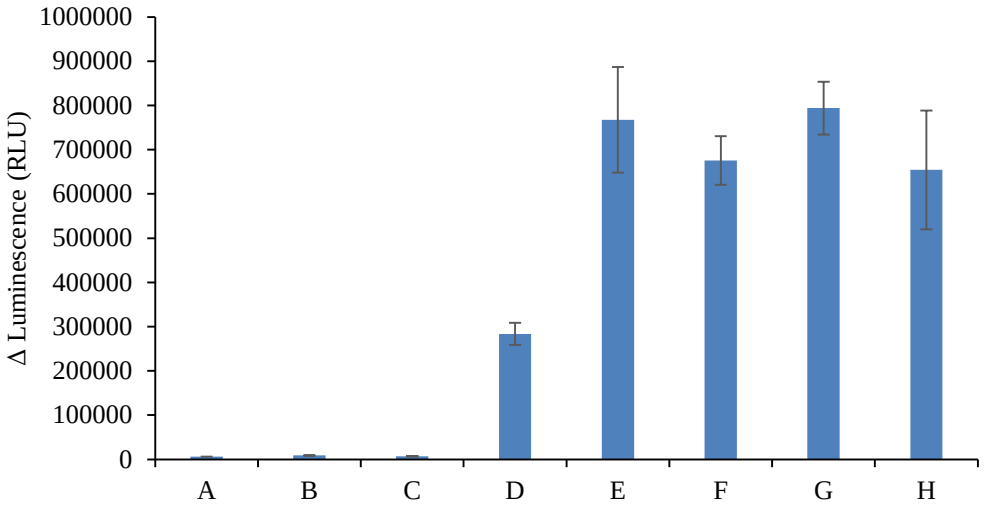


Figure 3.22 – Schematic of the NAD(P)H-Glo Kit (Promega) used to detect the NADH production from the ABA2 enzyme. The reductase substrate is a proprietary substrate provided within the kit. The luciferin detection reagent provided contains a luciferase enzyme and adenosine triphosphate (ATP).

It is possible that NAD^+ was present somehow in either the 9'-*cis*-neoxanthin extract, AtABA2 or ZmNCED1, but it is unlikely that the NAD^+ was carried through the purification of ABA2 or NCED, since it is bound only loosely by enzymes. It was noted in early experiments that there was some slight precipitation in the assays, which may have interfered with readings. However, this precipitation was eliminated through the use of cleaner preparations, meaning that it could not be the cause of the high background readings. In addition, ABA2 was further purified to eliminate other proteins which may interfere with the assays, but again this had no effect on the assay.



	A	B	C	D	E	F	G	H
Buffer	+	+	+	+	+	+	+	+
Catalase		+		+				
NAD^+			+	+	+	+	+	+
AtABA2					+	+	+	
9'-*cis*-neoxanthin					+		+	+
ZmNCED1					+	+		+

Figure 3.23 – Changes in luminescence seen for controls and full coupled assay using the NAD(P)H Glo kit measuring NADH production.

It is clear that there are inconsistencies across the assay which render the assay unreliable at this moment. Possible solutions include obtaining completely pure samples of ABA2 and NCED in order to totally remove competition from contaminating non-specifically bound proteins which may be inhibiting the assay. Low turnover of NCED may also be a problem and repeating the experiment with a high concentration of NCED to ensure complete turnover of 9'-*cis*-neoxanthin by NCED, thus overcoming any background, may be a possibility.

3.7 Conclusions

It is difficult to draw any firm conclusions on what features of the hydroxamic acids are required for inhibition of NCEDs (summarised in Figure 3.24). It is, however, interesting to note the time dependent inhibition observed with compounds D2 and D4 against ZmNCED1. It is entirely possible that the weak inhibition data seen against LeNCED1 is due to a lack of pre-incubation. This would in part explain why there is a discrepancy between *in*

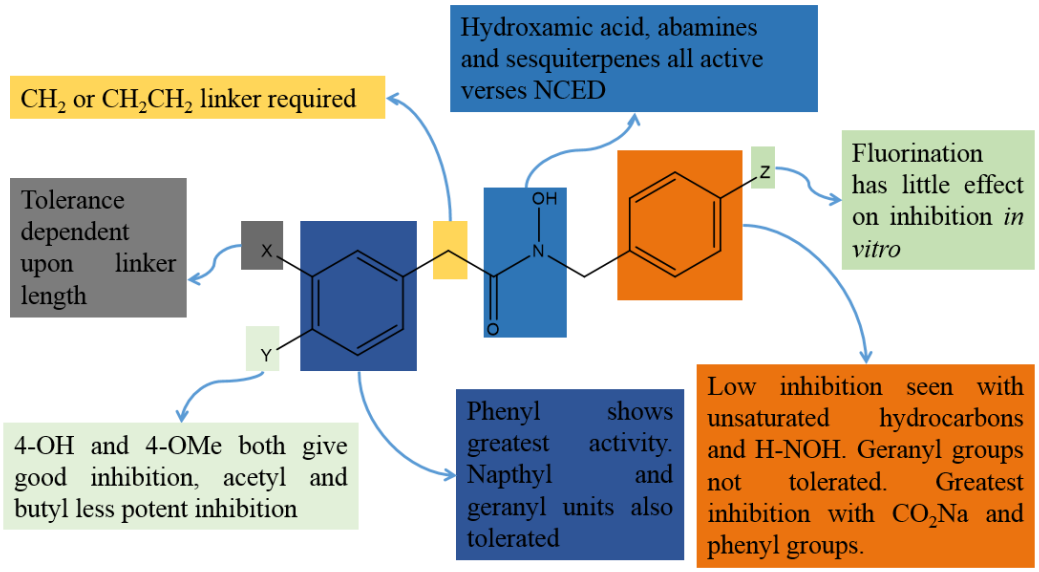


Figure 3.24 – Summary of the *in vitro* structure activity relationship data of hydroxamic acid compounds against cell free extract containing over produced LeNCED1.

planta data, which suggests inhibition of ABA mediated processes, and the *in vitro* data shown above. Poor inhibition by compounds abamine and abamineSG, which inhibit ABA mediated process *in planta*, may suggest an additional target is involved *in planta*.

The apparent lack of any difference in inhibition between the five NCEDs in *Z. mays* is understandable given the magnitude of the sequence similarity between the homologues. Given that the different NCEDs are expressed at different stages of plant development, in order to develop a commercial NCED inhibitor it would be necessary to develop a specific NCED inhibitor. However, since time dependent inhibition of ZmNCED1 has been shown, there could well be a difference in inhibition between the different ZmNCED homologs towards D2 and D4.

Finally, attempts to develop a method for the high throughput assay of NCED by coupling to the enzyme ABA2 were ultimately unsuccessful. Although the activity of each enzyme can be shown individually, upon addition of the two enzymes there is a significant amount of background, making it impossible to separate any actual reading from that background. Unfortunately, despite many experiments, the causes of the high background readings in the measurements remains unknown.

3.8 Future Work

Given the time dependent inhibition effect of hydroxamic acids *in vitro* towards Vp14, it would be of interest to assay the library of hydroxamic acids against ZmNCED1 and LeNCED1 and obtain IC_{50} values in order to discern structure activity relationships, which would aid in the development of selective NCED inhibitors.

Further development of new inhibitors is required to ascertain whether any of the structure activity features highlighted here hold true. A crystal structure with bound inhibitor would be invaluable for this endeavour. It would be interesting to investigate whether the hydroxyl groups and epoxide on 9-*cis*-neoxanthin are required for binding. In addition to this, it is known that NCED only accepts 9-*cis* substrates and developing an inhibitor which mimics this could be an interesting avenue of investigation. This may be why the longer chained inhibitors are more effective, as free σ bond rotation may allow the inhibitor to access conformations thus allowing it to mimic the *cis* double bond of the 9-*cis* substrates.

Experiments with *Oryza sativa* Dwarf27 have shown that D27 does not isomerise 9'-*cis*-neoxanthin to the all-*trans* isomer (see Chapter 4). Identifying the elusive neoxanthin isomerase, NCEI, would be of interest. A BLAST search reveals there are three homologs of D27 in *A. thaliana*, one of which may be the neoxanthin isomerase.

Finally, further work in developing a coupled assay would be beneficial, although potential problem is the low activity of the purified NCED. If the activity of the NCED can be increased, then it may be possible to discern a signal above the background of the reaction.

Chapter Four: Expression, Inhibition, Assays and Mechanistic Studies of *Oryza sativa* Dwarf27 (D27)

4.1 Introduction

The biosynthesis of strigolactone begins with all-*trans*-*β*-carotene, which is isomerised by an isomerase enzyme termed Dwarf27, or D27, to yield 9-*cis*-*β*-carotene (Figure 4.1). The isomerised *β*-carotene is then first cleaved by carotenoid cleavage dioxygenases 7 at the 9',10' position to give one equivalent each of *β*-ionone and 9-*cis*-*β*-apo-10'-carotenal. The 9-*cis*-*β*-apo-10'-carotenal is then cleaved by the second CCD in the pathway, CCD8, to produce a compound known as carlactone (Alder *at al*. 2012).

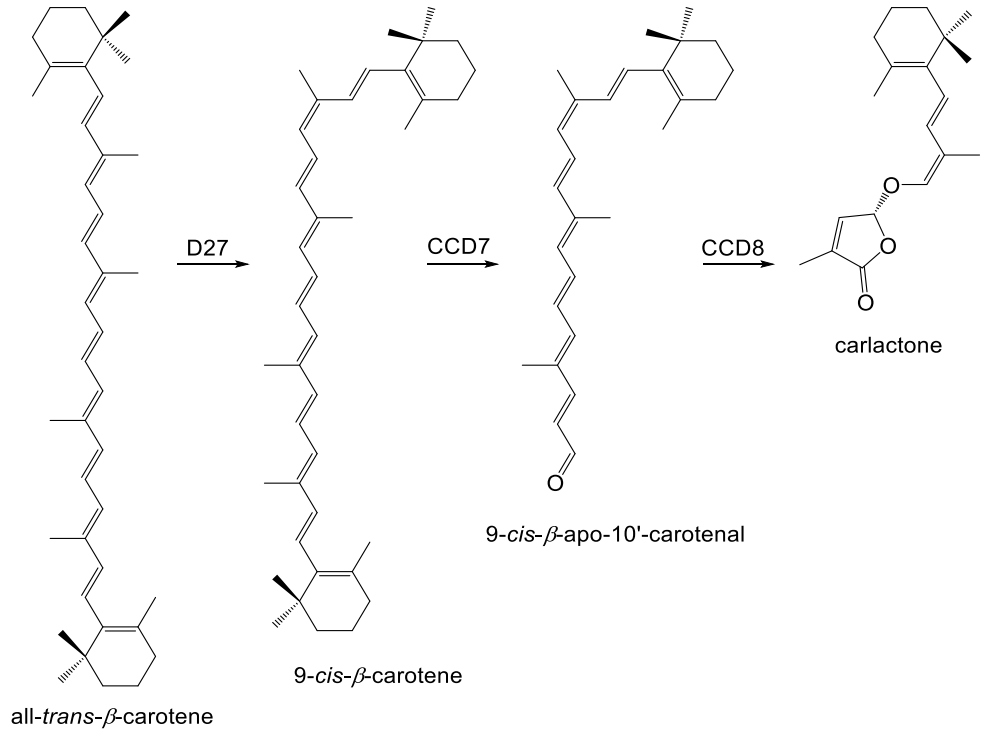


Figure 4.1 – Schematic of the biosynthetic pathway from all-*trans*-*β*-carotene to carlactone catalysed by D27, CCD7 and CCD8.

In planta, phenotypic effects on shoot branching have been observed with the application of hydroxamic acids, in particular D2, D4, D5 and D6, to the model organism *Arabidopsis thaliana* (Sergeant *et al*. 2009). Given that this class of hydroxamic acids has been shown to inhibit CCDs, it is likely the cause of this observed shoot branching phenotype

is the inhibition of either CCD7 or CCD8. Interestingly, the enzyme D27 is an iron containing enzyme (Lin *et al*. 2009). The role of iron is unknown, and possible isomerisation mechanisms involving a catalytic iron centre can be postulated (Chapter 1.7). The possibility of inhibition of D27 by hydroxamic acids cannot be ignored. Expression of D27 was therefore of interest for both generation of substrates for CCD7 and CCD8, and for enzyme inhibition assays.

The work detailed within this chapter describes the cloning of the enzyme D27, inhibition assays of hydroxamic acids towards D27 and mechanistic studies of D27. Cloning of D27 into the pGEX-4T-1 and pMAL-c5e vector and initial *in vitro* assays were performed by MChem project student Sophie Newgas.

4.2 Cloning and Expression of *O. sativa* Dwarf27

The cDNA of full length *Oryza sativa* Dwarf27 (herein termed OsD27.1) was obtained from the Genbank database and the synthetic DNA sequence was sequence optimised for expression in *E. coli* and purchased from Genscript (Lin *et al*. 2009) (See Appendix 10.4 for optimised sequence. Wild type accession number: FJ641055). The D27 genomic DNA was cloned from the pUC-57-Kan-OsD27.1 plasmid into a pET-151 expression vector via TOPO cloning.

Expression of N-His_6-OsD27.1 was induced with 1 mM IPTG and cultures were allowed to incubate overnight at 20º C. Following purification via nickel affinity chromatography there was no band corresponding to the N-His_6-OsD27.1 protein in the total soluble protein fraction following SDS-PAGE analysis. However, a band corresponding N-His_6-OsD27.1 was present in the total protein fraction (Figure 4.2). Attempts to refold the protein using 6 M guanidinium hydrochloride and dialysis were unsuccessful. Induction at different temperatures (16° C and 25° C) with different concentrations of IPTG (500 µM and 250 µM) also failed to produce soluble OsD27 protein.

Analysis of the sequence of OsD27 using the UniProt database reveals that the first 40 amino acids of D27 correspond to a potential chloroplast transit peptide. Consequently, the

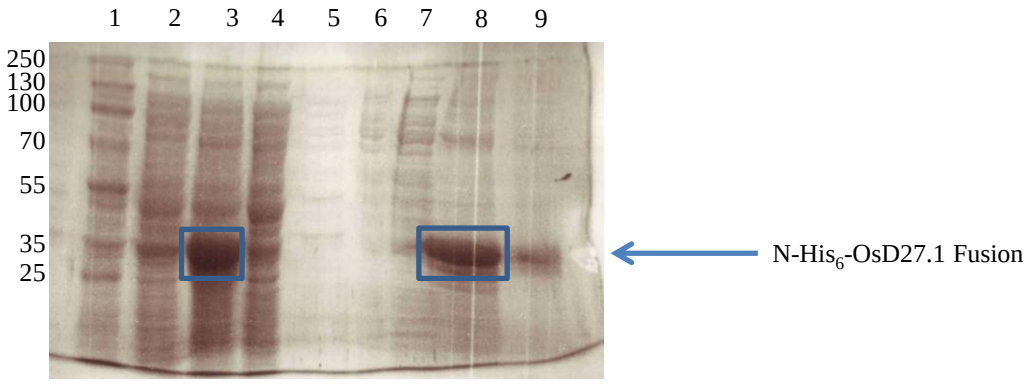


Figure 4.2 – 12% SDS-PAGE gel showing overproduced N-His_6-OsD27.1 in the insoluble total protein fraction (lane 3) from pET-151-OsD27.1 BL21 *E. coli*. 1 – Ladder; 2 – Non-induced control; 3 – Total protein; 4 – Total soluble protein from *E. coli* induced with IPTG (cell free extract); 5 – 6 M Guanidinium hydrochloride fraction; 6 – 5 µL Dialysed protein; 7 – 25 µL Dialysed protein; 8 – 20 µL Insoluble pellet following dialysis; 9 - 5 µL Insoluble pellet following dialysis. Molecular weights of ladder are in KDa. Molecular weight of N-His_6 tagged OsD27.1: 34.9 kDa (calculated from sequence).

first 120 base pairs were cloned out of OsD27 (OsD27ΔTP (termed D27.2)). OsD27.1 and OsD27.2 were subsequently cloned into the pMAL-c5e and the pGEX-4T-1 expression vectors as well as the pET-151 expression vector to give the maltose binding protein (MBP), glutathione-S-transferase (GST) and N-His_6 tagged recombinant proteins. The large soluble nature of the GST and MBP tags was predicted to aid the solubility of D27. The expression of the MBP, GST and N-His_6 tagged OsD27.1 and OsD27.2 was tested in the presence and absence of the pBB541 plasmid. The pBB541 plasmid was obtained from Addgene (de Marco 2007) and contains the genes *groES* and *groEL*, coding for the GroES and GroEL proteins. GroES and GroEL are chaparones from *E. coli* which, when overproduced in conjunction with a protein, aid the correct folding of that particular protein. This helps to ensure solubility and prevent aggregation of that protein.

A summary of the results from the expression tests is presented in Table 4.1. In all cases protein over production was induced with 1 mM IPTG and cultures were incubated for 16 hours at 20° C. All cloning experiments were performed successfully and results confirmed by sequencing. Only the MBP-OsD27.1 protein construct failed to express, the reasons for which are unknown. Removal of the transit peptide failed to improve the solubility of N-His_6

Construct	N-His$_6$-OsD27.1	N-His$_6$-OsD27.1	N-His$_6$-OsD27.2	N-His$_6$-OsD27.2	GST-OsD27.1*	GST-OsD27.1	MBP-OsD27.1*
(+) pBB541 Plasmid?	(-)	(+)	(-)	(+)	(-)	(+)	(-)
Expression?	Y	Y	Y	Y	Y	Y	N
Soluble?	N	Y	N	Y	Y	Y	-
Active?	-	N	-	Y	Y	Y	-

Table 4.1 – Summary of expression constructs tested for expression of the OsD27 protein. Constructs marked with '*' were prepared by Sophie Newgas. OsD27.2 lacks the 120 amino acid transit peptide.

tagged OsD27. Expression of either N-His$_6$-OsD27.1 or N-His$_6$-D27.2 did not result in soluble protein unless the GroES/EL chaperones were co-expressed. GST-OsD27.1 was also found to be soluble and active. In all cases, the level of soluble protein was low (Figure 4.3). For *in vivo* work, the N-His$_6$ tagged OsD27.1 protein was used, whilst for *in vitro* work, either GST tagged OsD27.1 in the presence of GroES/EL or N-His$_6$ tagged OsD27.2 in the presence of GroES/EL was used.

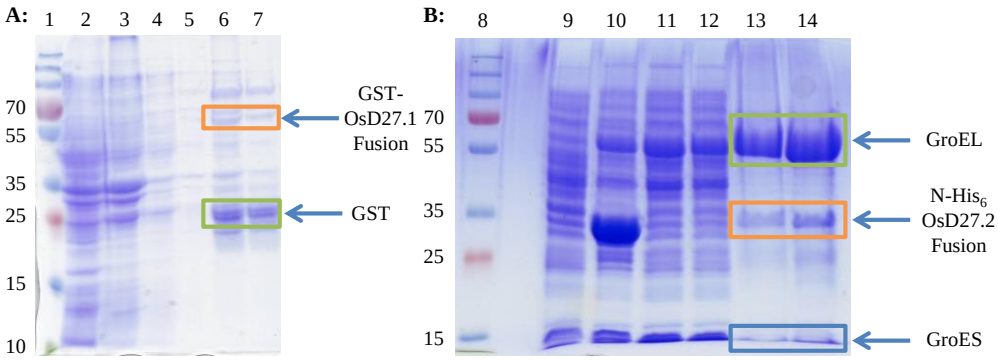


Figure 4.3 – **A:** 10% SDS-PAGE gel showing overproduced GST-OsD27.1 from pGEX-4T-1-OsD27.1 BL21 *E. coli* following Coomassie staining. 1 – Ladder; 2 – Total soluble protein; 3 – Unbound protein; 4 – Elution fraction 1; 5 – Elution fraction 2; 6 – Elution fraction 3; 7 – Elution fraction 4; 8 – Elution fraction 5. **B:** 10% SDS-PAGE gel showing overproduced N-His$_6$-OsD27.2 from pET-151-OsD27.2 BL21 *E. coli* following Coomassie staining. 8 – Ladder; 9 – Non-induced control; 10 – Total protein; 11 – Total soluble protein; 12 – Flow through; 13 – Elution fraction; 14 – Concentrated elution fraction. Molecular weights of ladder are in kDa. Molecular weight of GST tagged OsD27.2: 57.1 kDa. Molecular weight of N-His$_6$ tagged OsD27.2: 30.4 kDa. Mass spectrometry would be required to definitively confirm the presence of the OsD27 protein.

4.3 Identification of *β*-Carotene Isomers via HPLC

Isomers of *β*-carotene can be identified via UV-visible spectroscopy due to the differences in the λ_{MAX} values which result from isomerisation of the double bond backbone. Since the λ_{MAX} values are overlapping, the isomerisation reaction cannot be assayed continuously. Table 4.2 summarises λ_{MAX} values reported in the literature for 9-*cis*-*β*-carotene and all-*trans*-*β*-carotene.

In order to differentiate between isomers of *β*-carotene, a hydrophobic C_{30} reversed phase HPLC column is required (Figure 4.4). The high hydrophobicity of the C_{30} column

	Emenhiser *et al.* 1995	Tsukida *et al.* 1982	Alder *et al.* 2012	Experimentally observed
Absorbance maxima for all-*trans* (nm)	450	450	453	452
Absorbance maxima for 9-*cis* (nm)	444	445	447	448

Table 4.2 – Comparison of λ_{MAX} values in the literature for all-*trans*-*β*-carotene and 9-*cis*-*β*-carotene to those obtained experimentally in this project.

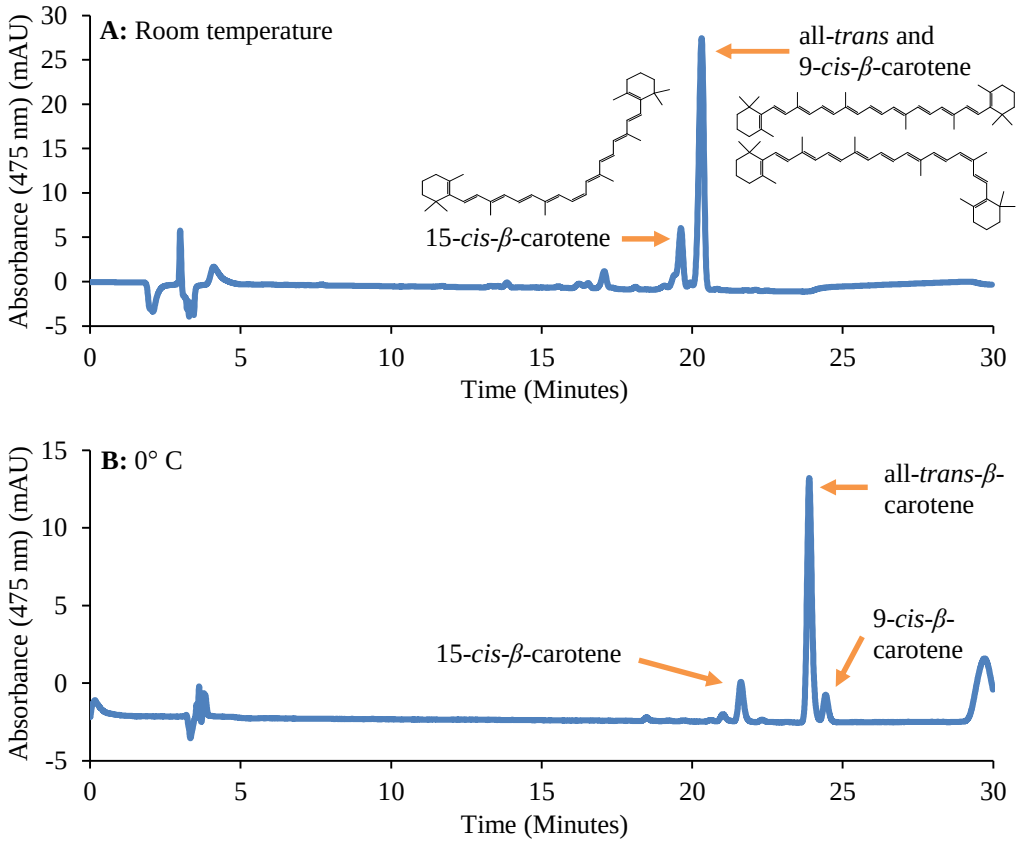


Figure 4.4 – Comparison of HPLC chromatograms at room temperature and 4° C for the separation and analysis of *β*-carotene and its isomers.

allows isomers of *β*-carotene to be separated using a suitable solvent system (Sharpless *et al*. 1996). It was noticed in order for complete separation of peaks, it was essential to operate the column at a temperature of 0º C. When the column is operated at room temperature, peaks with close retention times co-elute, thus preventing separation of certain isomers. However, at a lower temperature, the retention and interaction of the isomers with the column is increased, allowing separation of the isomers (Figure 4.4).

Analysis of the peaks at 23.9 minutes and 24.5 minutes showed that the first peak had λ_{MAX} values at 452 nm and 476 nm (Figure 4.5). The second peak on the other hand had λ_{MAX} values at 448 nm and 474 nm (Figure 4.5). These values are in agreement with those reported in the literature for all-*trans*-*β*-carotene and 9-*cis*-*β*-carotene respectively (Tsukida *et al*. 1982, Emenhiser *et al*. 1995, Alder *et al*. 2012). In addition to this, *cis*-isomers of *β*-carotene have a

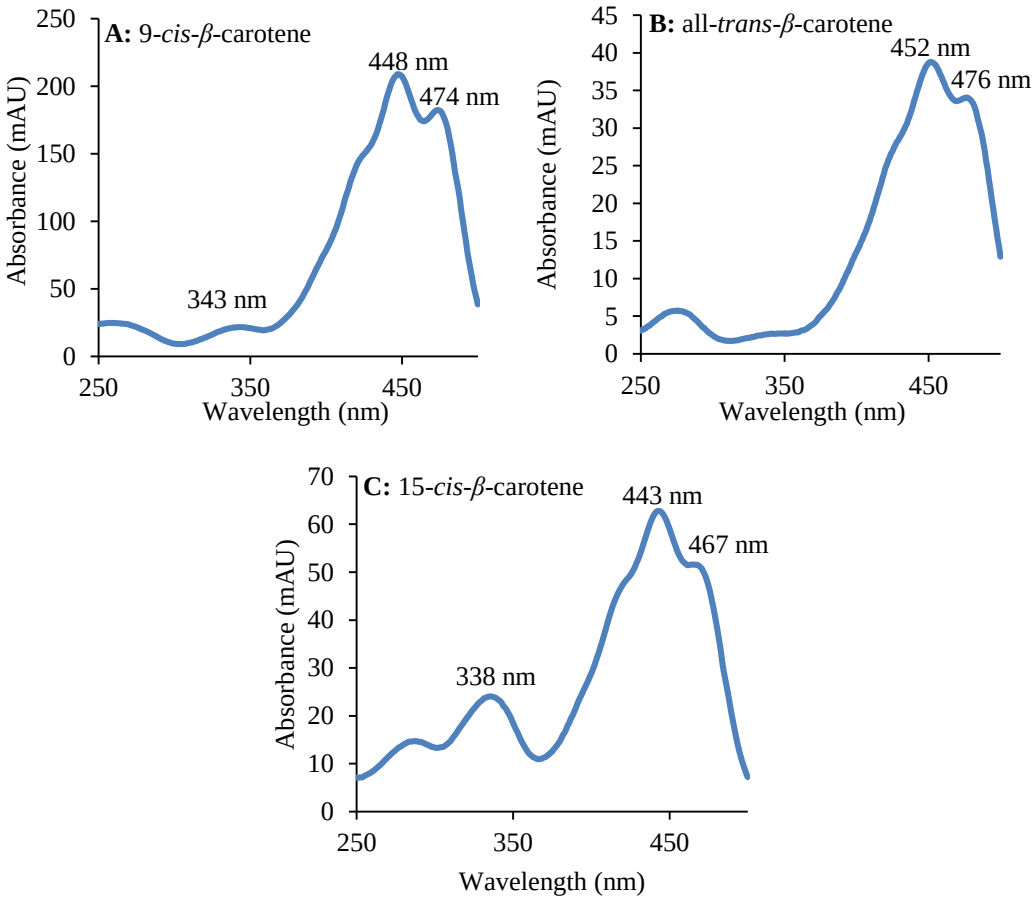


Figure 4.5 – Comparison of the UV absorption spectra of 9-*cis*-*β*-carotene (**A**, left), all-*trans*-*β*-carotene (**B**, right) and 15-*cis*-*β*-carotene (**C**, bottom).

cis peak between 320 and 350 nm, 343 nm in the case of 9-*cis*-*β*-carotene. The wavelength of this *cis* peak can be used to identify *β*-carotene isomers (Tsukida *et al*. 1982). Using these techniques, the peak at 23.9 minutes was identified as belonging to the all-*trans*-isomer and the peak at 24.5 minutes was assigned to the 9-*cis* isomer. The peak at 21.6 minutes was identified as belonging to the 15-*cis* isomer of *β*-carotene via its UV absorbance profile (Figure 4.5). It should be noted, however, that comparison to synthetic authentic standards would be required for unambiguous assignment of the *β*-carotene isomers.

4.4 *In vivo* assays of *O. sativa* Dwarf27

For assays of OsD27, *β*-carotene was extracted using 6 M pyridinium acetate pH 5.5, as discussed in Section 5.3. Analysis was performed using a C_{30} reversed phase column as described above. Due to the low solubility of OsD27.1 and OsD27.2, *in vivo* assays of OsD27 were attempted using a *β*-carotene accumulating strain of *E. coli*, pAC-BETA BL21 *E. coli*, which contains the pAC-BETA plasmid (obtained from Dr Andrew Thompson (Cranfield University)), which produces *β*-carotene when induced with 100 µM IPTG. Chemically competent pAC-BETA BL21 *E. coli* was prepared and transformed with the pET-151-OsD27.1 plasmid. Although the N-His_6-OsD27.1 was insoluble *in vitro*, it was suspected that *in vivo* a large proportion of the protein would remain soluble. The two strains of *E. coli*, pAC-BETA BL21 *E. coli* and pAC-BETA pET-151-OsD27.1 BL21 *E. coli* were grown to OD_{600} = 0.6 AU and induced with 100 µM IPTG.

After incubation for 16 hours at 20º C, the cells were collected by centrifugation and lysed via the action of lysozyme and sonication. Reaction products and unreacted substrate were extracted with 6 M pyridinium acetate and *n*-butanol extraction. Analysis showed that in the presence of pET-151-OsD27.1 the ratio of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene was smaller than with respect to the control (pAC-BETA in the absence of D27) (Figure 4.6) indicating that *in vivo* D27 isomerises all-*trans*-*β*-carotene to the 9-*cis* isomer. *In vivo* isomerisation was additionally performed with pGEX-4T-1-OsD27.1, demonstrating the activity of the construct.

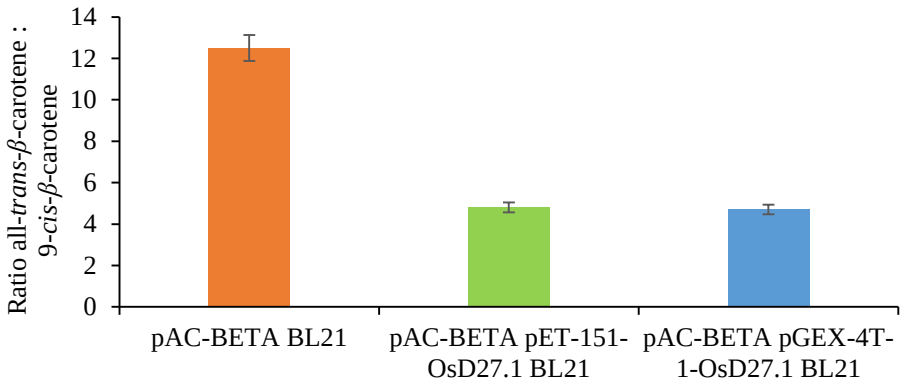


Figure 4.6 – Comparison of the ratio of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene produced by *E. coli* expressing the pAC-BETA *β*-carotene producing plasmid in the presence and absence of a plasmid expressing the D27 *β*-carotene isomerase. Lower ratio of all-*trans* to 9-*cis*-*β*-carotene indicates there is proportionally more 9-*cis*-*β*-carotene and hence isomerisation has occurred.

In order to test whether the hydroxamic acids were able to inhibit D27, *in vivo* cultures of pAC-BETA BL21 *E. coli* and pAC-BETA pET-151-OsD27.1 BL21 *E. coli* were prepared as before in the presence of 100 µM D4 (Figure 4.7). Unfortunately, on analysis it was found there was no difference in the ratio of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene in the presence and absence of D4. This is by no means conclusive proof that hydroxamic acids do not inhibit OsD27, given that it is possible that D4 is simply not taken up into the cell by *E. coli*. On the basis that the *in vivo* data showed activity of the pGEX-4T-1-OsD27.1, *in vitro* experiments were attempted.

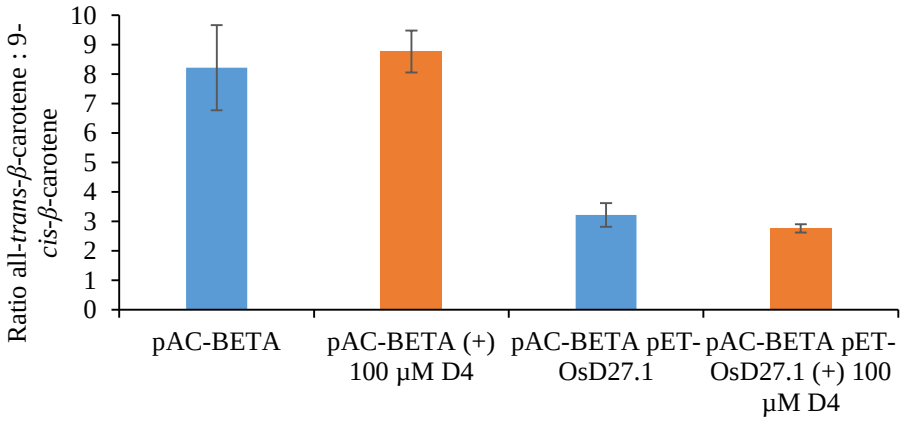


Figure 4.7 – Comparison of the ratio of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene produced by *E. coli* expressing the pAC-BETA *β*-carotene producing plasmid in the presence and absence of a plasmid expressing the D27 *β*-carotene isomerase and in the presence and absence of 100 µM D4.

4.5 *In vitro* assays of *O. sativa* Dwarf27

Initial *in vitro* assays with GST-D27.1 were performed by Sophie Newgas. In these experiments, partially purified GST-OsD27.1 (50 µg total protein) was used with commercial *β*-carotene (99% all-*trans* isomer, 40 µM). Assays were incubated at 20° C in the dark for 2 hours. Following analysis by HPLC it could be seen that there was only a very small change in the ratio of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene with respect to the control, indicating that little or no reaction had taken place (Figure 4.8). Moreover, it was observed that this small change in the ratio of *β*-carotene isomers could be due to non-specific isomerisation. When a geometrically pure solution of all-*trans*-*β*-carotene was obtained following purification via HPLC and carried through the assay in absence of OsD27.1, then reanalysed, new peaks corresponding to the 15-*cis* and 9-*cis* isomers were present (Figure 4.9).

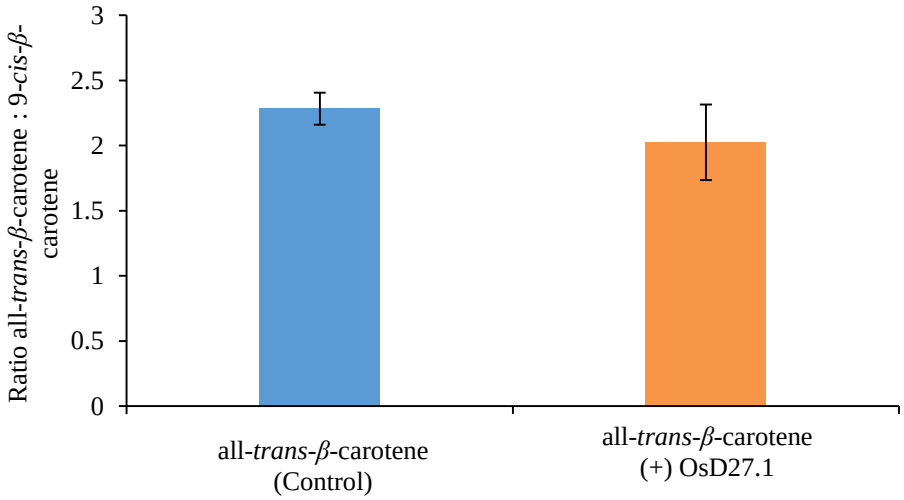


Figure 4.8 – Comparison of the ratio of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene in the presence and absence of OsD27.1.

Isomerisation of a solution of all-*trans*-*β*-carotene by iodine under UV light results in a mixture of different geometrical isomers in equilibrium (See chapter 5.3). Analysis of these isomers via HPLC gives an equilibrium constant (K_{eq}), defined by Equation 4.1, of 0.53 for the equilibrium of all-*trans*-*β*-carotene and 9-*cis*-*β*-carotene. D27 is catalysing a homologous

$$K_{eq} = \frac{[9\text{-}cis\text{-}\beta\text{-carotene}]}{[\text{all-}trans\text{-}\beta\text{-carotene}]}$$

Equation 4.1 – Equation for the calculation of the equilibrium constant for the isomerisation of 9-*cis*-*β*-carotene and all-*trans*-*β*-carotene.

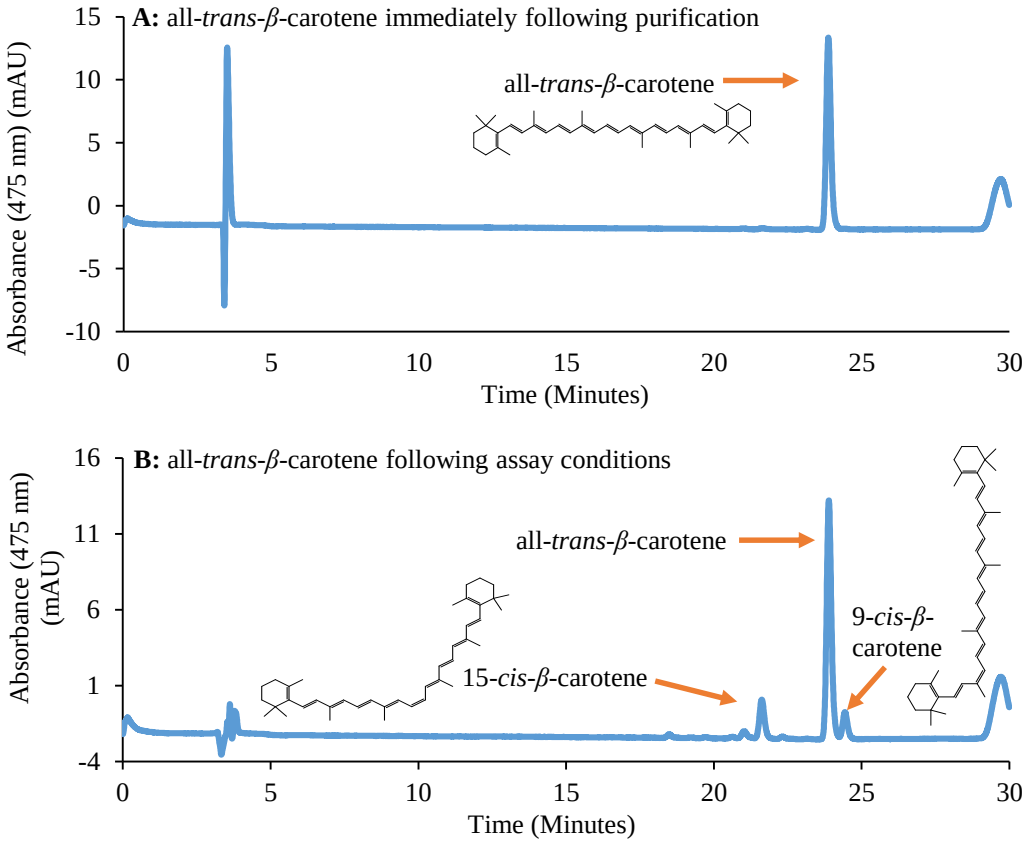


Figure 4.9 – Comparison of purified all-*trans*-*β*-carotene immediately following purification via HPLC (**A**, top), and following exposure to assay conditions (**B**, bottom). Non-specific, non-enzyme catalysed isomerisation occurs as a background reaction.

reaction to that of iodine with UV light. It would be expected that following incubation of all-*trans*-*β*-carotene with D27, the *β*-carotene solution should reach a similar equilibrium. However, this was apparently not the case.

Due to the low activity of OsD27 and the inaccuracies of observing the isomerisation reaction due to competing non-specific isomerisation, and to confirm that D27 was active, D27 was coupled to the second enzyme on the strigolactone biosynthesis pathway, CCD7 (assays of AtCCD7.2 are discussed in Chapter 5). Incubation of all-*trans*-*β*-carotene with CCD7 failed to produce a peak corresponding to the 9-*cis*-*β*-apo-10-caroteal product of the CCD7 reaction (Figure 4.10). However, when D27 is incubated with AtCCD7.2 and all-*trans*-*β*-carotene, a peak correspond to the 9-*cis*-*β*-apo-10'-carotenal, the product of the CCD7 reaction, becomes apparent (Figure 4.10 and 4.11). This confirms that the D27 is active and

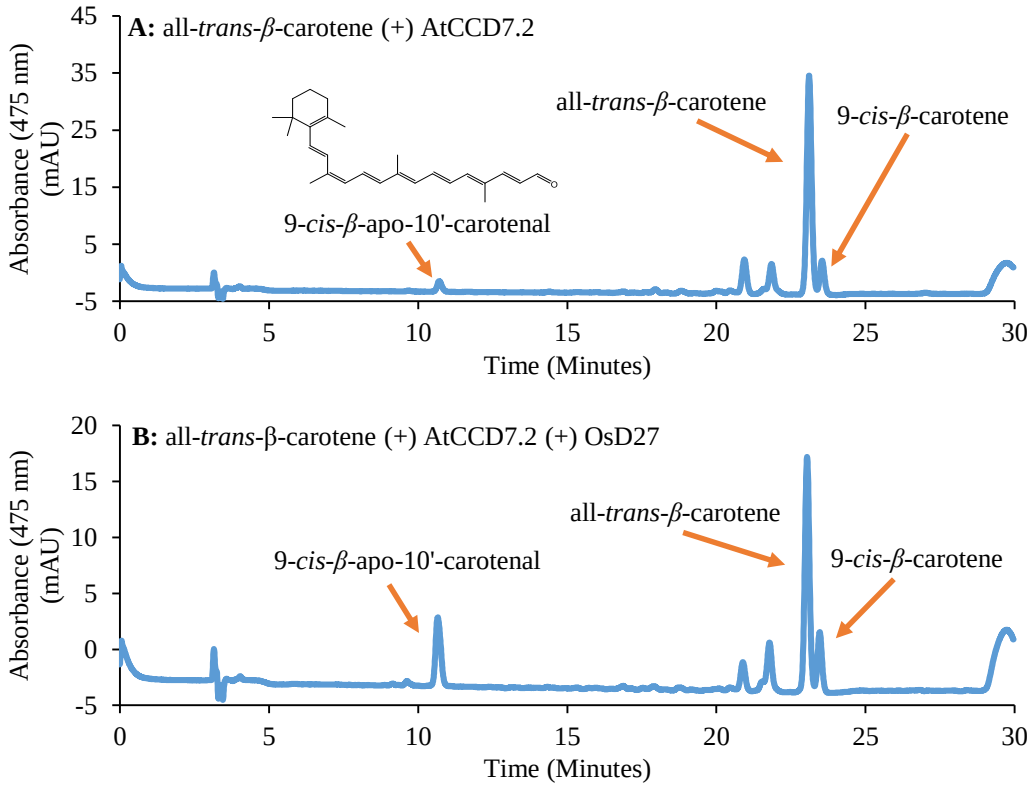


Figure 4.10 – HPLC chromatogram of products from the incubation of all-*trans*-*β*-carotene with AtCCD7.2 (**A**, top) and of all-*trans*-*β*-carotene with OsD27 and AtCCD7.2 (**B**, bottom).

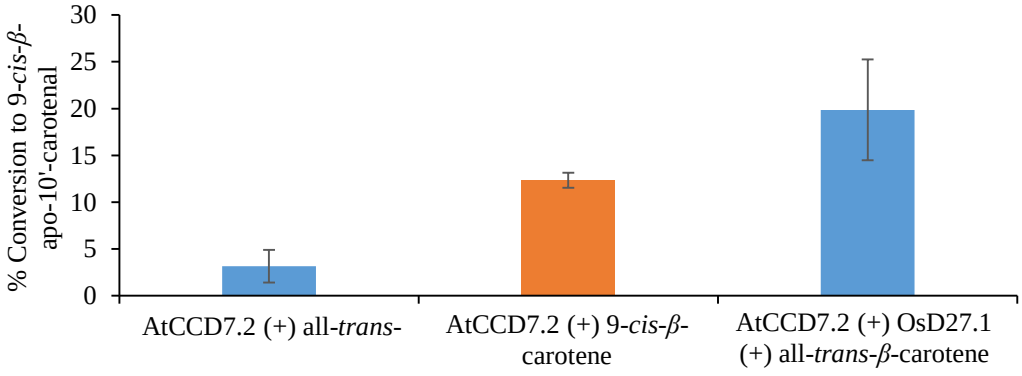


Figure 4.11 – Percentage conversion of 9-*cis*-*β*-carotene to 9-*cis*-*β*-apo-10'-carotenal. In the absence of the isomerase OsD27.1, there is less than 5% conversion of all-*trans*-*β*-carotene by AtCCD7.2. However, when AtCCD7.2 is incubated with all-*trans*-*β*-carotene and OsD27.1 the level of conversion increases above that observed when AtCCD7.2 is incubated with 9-*cis*-*β*-carotene. 9-*cis*-*β*-apo-10'-carotenal was identified by mass spectrometry (see Chapter 5).

does catalyse the isomerisation of all-*trans*-*β*-carotene to the 9-*cis* isomer. However, it raises the question as to why D27 does not appear to catalyse the isomerisation on its own?

It was investigated as to whether D27 could catalyse the reverse reaction, the isomerisation of 9-*cis*-*β*-carotene to all-*trans*-*β*-carotene. When D27 is incubated with a

solution of *β*-carotene containing predominantly the 9-*cis*-isomer, and the products analysed it can be seen that there is a significant increase in the levels of all-*trans*-*β*-carotene present (Figure 4.12). The amount of all-*trans*-*β*-carotene increases until an equilibrium between the *β*-carotene isomers is reached (Figure 4.13). Following incubation of D27 with 9-*cis*-*β*-carotene, the K_{eq} of the isomers equals 0.49, similar to that of the non-enzyme catalysed iodine isomerisation.

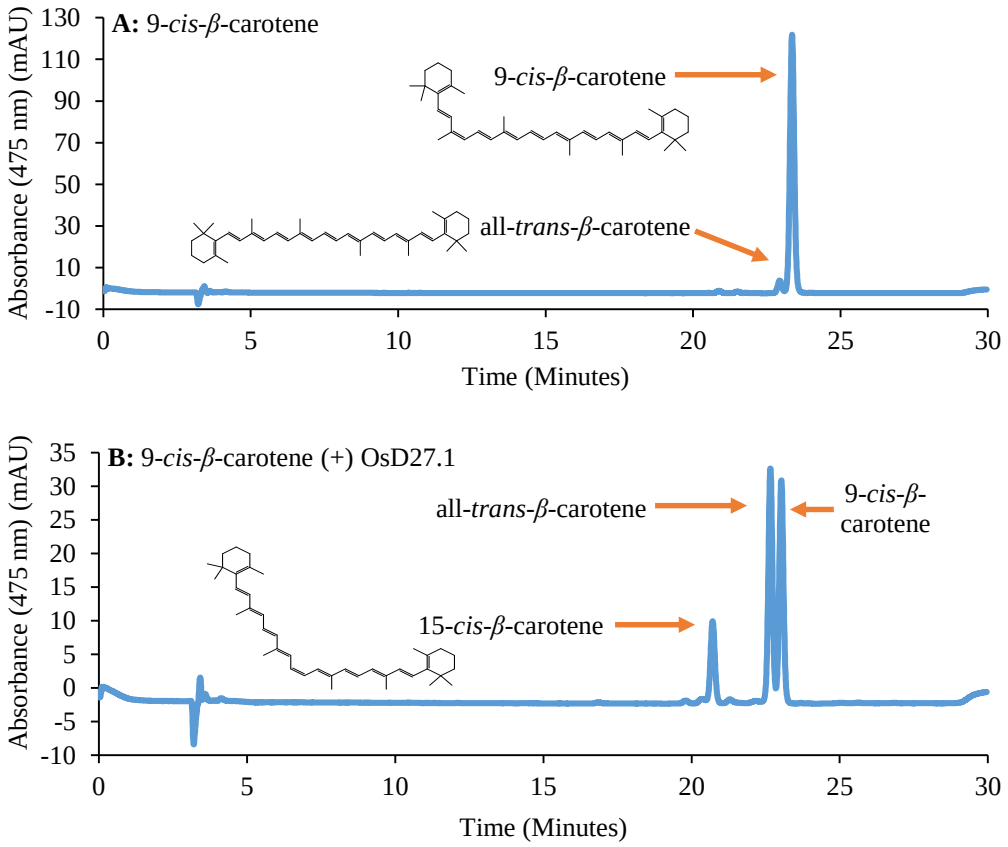


Figure 4.12 – Incubation of 9-*cis*-*β*-carotene in the absence (**A**) and presence (**B**) of OsD27.1. Following incubation of 9-*cis*-*β*-carotene with OsD27.1 the 15-*cis* and all-*trans* isomers of *β*-carotene are visible, indicating that the reverse isomerisation reaction is favoured and the enzyme is not selective for solely the all-*trans* and 9-*cis* isomers of *β*-carotene.

Consequentially, the question becomes why is D27 not catalysing the forward reaction to completion? There are two possible solutions to this question. The first, inactivation of the enzyme, is unlikely, given that the D27 has been shown to be active. The second is product inhibition. Given that the 9-*cis* isomer is thermodynamically higher in energy than the all-*trans* isomer due to the favourable overlap of the π orbitals in the all-*trans* state and decreased steric interactions (Niedzwiedzki *et al*. 2009), it is likely that the enzyme binds the 9-*cis* isomer

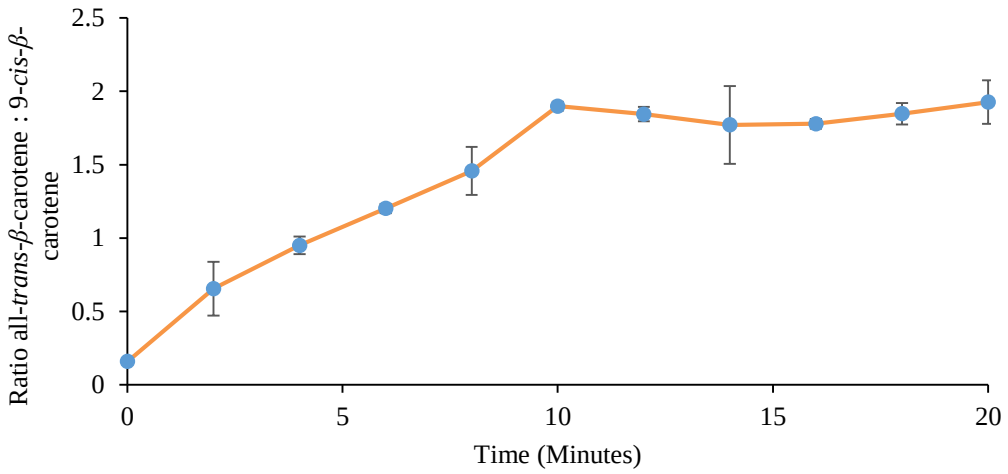


Figure 4.13 – Time course experiment of OsD27 in the presence of 9-*cis*-*β*-carotene. The reaction proceeds until an equilibrium is reached between the isomers of *β*-carotene.

more tightly than the all-*trans* isomer. If so, the forward reaction forms a higher affinity product, which might lead to product inhibition. In the presence of CCD7, the 9-*cis*-*β*-carotene product would be readily removed, thus ensuring that the effective concentration of 9-*cis*-*β*-carotene in the reaction solution remains low, preventing product inhibition of D27. In the reverse reaction (9-*cis* to all-*trans*), the formation of lower affinity all-*trans*-*β*-carotene would not cause product inhibition.

A K_M of 0.26 µM was determined for OsD27 for 9-*cis*-*β*-carotene at pH 6.4 at 25° C, a very low K_M value. Additionally, a k_{cat} of 34 s^{-1} was calculated for the same reaction under the same conditions. However, due to poor efficiency of the forward isomerisation reaction, it was not possible to calculate either K_M or k_{cat} for OsD27 with all-*trans*-*β*-carotene. Despite this, the low K_M value for the reverse reaction implies that D27 does bind 9-*cis*-*β*-carotene with a high affinity, supporting the hypothesis of product inhibition by 9-*cis*-*β*-carotene in the forward reaction. Consequentially, this would imply that the K_M for the forward reaction, the isomerisation of all-*trans*-*β*-carotene, is higher than that of the reverse reaction. As such, since the K_{eq} of the reaction can also be defined by Equation 4.2, it must therefore follow that the k_{cat} of the forward reaction must be of a similar magnitude to that of the K_M, since the product of k_{cat} / K_M for the forward reaction must equal roughly 1 to give a K_{eq} of 0.5. This is supported by the observation that when D27 is coupled to CCD7 with all-*trans*-*β*-carotene there are

$$K_{eq} = \frac{\left(\frac{k_{cat}}{K_M}\right)_{Forward}}{\left(\frac{k_{cat}}{K_M}\right)_{Reverse}}$$

Equation 4.2 – Equation for the calculation of the equilibrium constant for the isomerisation of 9-*cis*-*β*-carotene and all-*trans*-*β*-carotene via K_M and k_{cat} values.

higher levels of 9-*cis*-*β*-apo-10'-carotenal formation than when CCD7 is incubated with 9-*cis*-*β*-carotene alone.

The 15-*cis* isomer can also be identified following incubation of 9-*cis*-*β*-carotene with D27 (Figure 4.12). The presence of the 15-*cis* isomer suggests that D27 is not 100% regiospecific, which could indicate that the isomerisation occurs through either a radical or cation mediated process, similar to the RPE65 isomerase, whereby loss of double bond character allows free rotation of the polyene backbone. An isomerisation reaction which utilises a covalent intermediate is unlikely to produce side products. Moreover it suggest that more than one bond can be isomerised by D27. However, the caveat with this data is that D27 catalysed isomerisation is competing with nonspecific thermal and UV catalysed isomerisation, and it is difficult to distinguish between the two using these assay methods.

Given the propensity of *β*-carotene to favour the all-*trans* isomer and that *in vitro* D27 activity is low, it is likely that *in vivo*, CCD7 and D27 are localised together. This would allow CCD7 to immediately cleave any 9-*cis*-*β*-carotene produced by D27. Removal of 9-*cis*-*β*-carotene would then prevent substrate inhibition of D27 and drive D27 to continue to isomerise all-*trans*-*β*-carotene to maintain the equilibrium of isomers.

4.6 Mechanistic Studies *O. sativa* Dwarf27

During these investigations, several questions regarding the mechanism by which D27 isomerises *β*-carotene became apparent, primarily by what mechanism does D27 isomerise *β*-carotene? D27 does not share any sequence similarity to the all-*trans*-retinyl ester isomerase, RPE65. In fact, analysis of D27 shows it is not related to any protein of known function. What is more, the isomerisation of retinol by RPE65 is dependent upon the ester moiety of the all-*trans*-retinyl ester, which is cleaved by the Lewis acidity of iron to create a substrate cation,

which can then isomerise (Kiser *et al*. 2012). Since *β*-carotene would be unable to interact with iron in a similar way with D27, this mechanism seems unlikely. As such, it is likely isomerisation occurs through a novel mechanism. For these assays, the isomerisation of 9-*cis*-*β*-carotene to all-*trans*-*β*-carotene was observed due to reasons outlined in Section 4.5.

A pH rate profile of D27 was performed to determine if any acid/base catalysis occurs in the mechanism of D27. Incubation of D27 with 9-*cis*-*β*-carotene at different pH revealed little variation in the rate of isomerisation of 9-*cis*-*β*-carotene to all-*trans*-*β*-carotene versus pH (Figure 4.14). As such, it is unlikely that there is any acid-base catalysis in the isomerisation of D27. This would therefore suggest a mechanism of isomerisation that is mediated by electron transfer to or from the carotenoid substrate. Removal or addition of an electron would create either a substrate radical cation or anion (Figure 4.17). This would break the conjugated double bond system and allow free rotation of the *β*-carotene. This hypothesis is supported by the observation that D27 is not regiospecific, producing isomers of *β*-carotene other than 9-*cis* or all-*trans*. The production of the 15-*cis* isomer can be clearly seen following incubation of 9-*cis*-*β*-carotene with OsD27 (Figure 4.12). A mechanism involving a covalent intermediate, similar to that suggested by Deigner *et al*. for the isomerisation of all-*trans*-retinol by RPE65 (Chapter 1.7, Figure 1.24), would be significantly less likely to produce side products (Deigner *et al*. 1989).

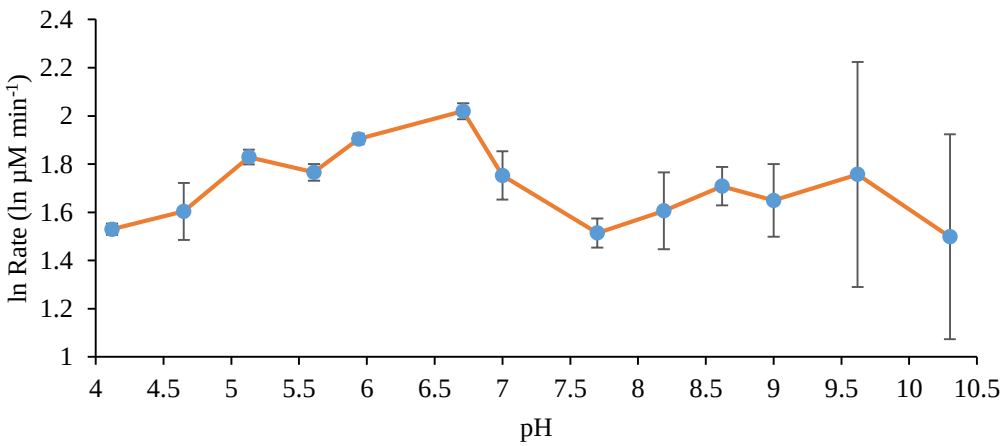


Figure 4.14 – pH rate profile of GST-OsD27.1, showing no significant variation in the rate of the D27 catalysed isomerisation of 9-*cis*-*β*-carotene to all-*trans*-*β*-carotene.

As discussed in Chapter 1.7, isomerisation of *β*-carotene could proceed through a mechanism whereby oxygen is used to create a substrate cation or radical (Figure 1.24). Formation of this intermediate results in a loss of double character, which allows free rotation and subsequent isomerisation (Harrison & Bugg 2013). In order to investigate whether D27 proceeds through this oxygen dependent mechanism, the isomerisation reaction was performed under an atmosphere of nitrogen gas following thorough degassing of the enzyme and buffers. Subsequent analysis of the reaction products revealed that the reaction occurred identically as under aerobic conditions (Figure 4.15). Moreover, a mechanism involving dioxygen might also be expected to produce the cleavage product, 9-*cis*-*β*-apo-10'-carotenal, in addition to the isomerised *β*-carotene. However, no such cleavage products are observed. Consequentially, these observations seem to rule out this mechanism.

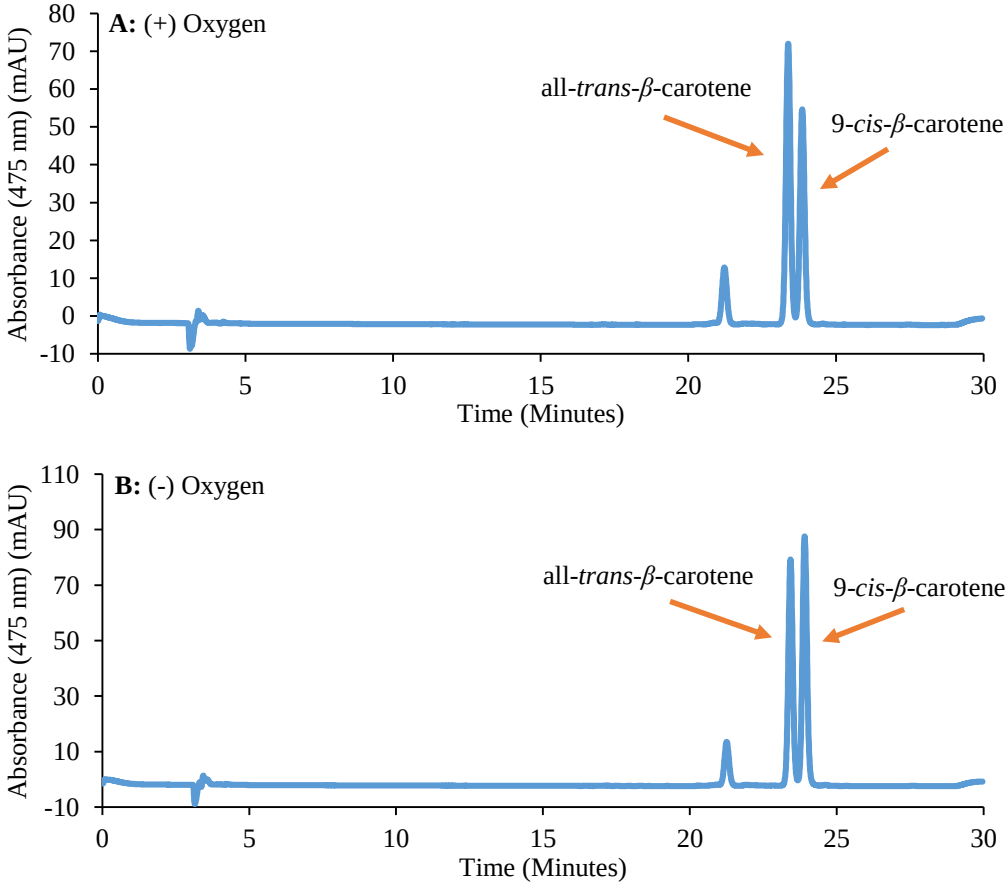


Figure 4.15 – HPLC chromatogram following incubation of 9-*cis*-*β*-carotene with OsD27 in the presence of oxygen (**A**, top) and absence of oxygen (**B**, bottom).

The enzyme prolycopene isomerase (CRTISO) is an isomerase required for the isomerisation of 7,9/9',7'-tetra-*cis*-lycopene to all-*trans*-lycopene (Isaacson *et al*. 2004). Although the mechanism of CRTISO is also unknown, experiments by Isaacson *et al*. have shown that purified CRTISO is inactive *in vitro*. CRTISO activity is only present upon the addition of bacterial cell free extract. Further experiments suggested CRTISO may require the presence of the respiratory transport chain proteins complex I (NADH ubiquinone oxidoreductase) and complex II (succinate ubiquinone oxidoreductase), indicating that isomerisation of 7,9/9',7'-tetra-*cis*-lycopene is a redox process, obtaining electrons from complex I before passing them to complex II (Isaacson *et al*. 2004). However, there is no direct evidence for this mechanism. Moreover, this raises the question of how does CRTISO mediate this electron transfer? CRTISO has been shown to contain a cofactor binding domain and as such, other possible mechanisms of isomerisation could be postulated, involving electron transfer to or from a redox active cofactor such as FAD(H) or NAD(P)(H). The resulting carotenoid cation, anion or radical would then be able to isomerise due to the loss of double bond character (Figure 4.16).

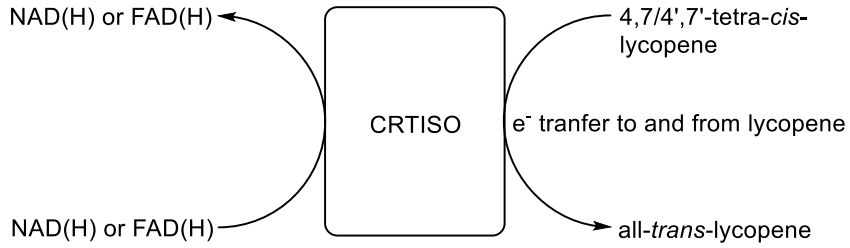


Figure 4.16 – Schematic of possible mechanism for the isomerisation of 4,7/4',7'-tetra-*cis*-lycopene to all-*trans*-lycopene by CRTISO.

Therefore, although sequence alignment showed that D27 is not related to CRTISO (data not shown) we therefore considered the possibility that D27 may contain a redox cofactor such as an iron sulfur cluster. It was reported by Lin *et al*. that D27 contains 1.7 moles of iron per mole of protein (Lin *et al*. 2009). If it assumed that there is in fact actually two moles of iron per mole of protein, then two possibilities exist. Either a di-nuclear [FeFe] centre or an [2Fe2S] cluster. Di-iron [FeFe] centres are commonly found in hydrogenase enzymes and are linked to an electron donor, such as an iron sulfur cluster or ferrodoxin (Ferraro *et al*. 2005).

It seems unlikely that D27 contains an [FeFe] centre for two reasons. Firstly, D27 is not dependent on the presence of a second protein, such as a ferrodoxin, and secondly, if D27 contained an iron sulfur cluster in addition to an [FeFe] centre, then the ratio of iron to protein would be higher than two to one. As such, the presence of a [2Fe2S] centre seems possible, performing a reaction similar to that shown in Figure 4.17, where the iron sulfur cluster mediates either the addition or abstraction of an electron to or from the carotenoid, resulting in a radical anion or cation which can then isomerise.

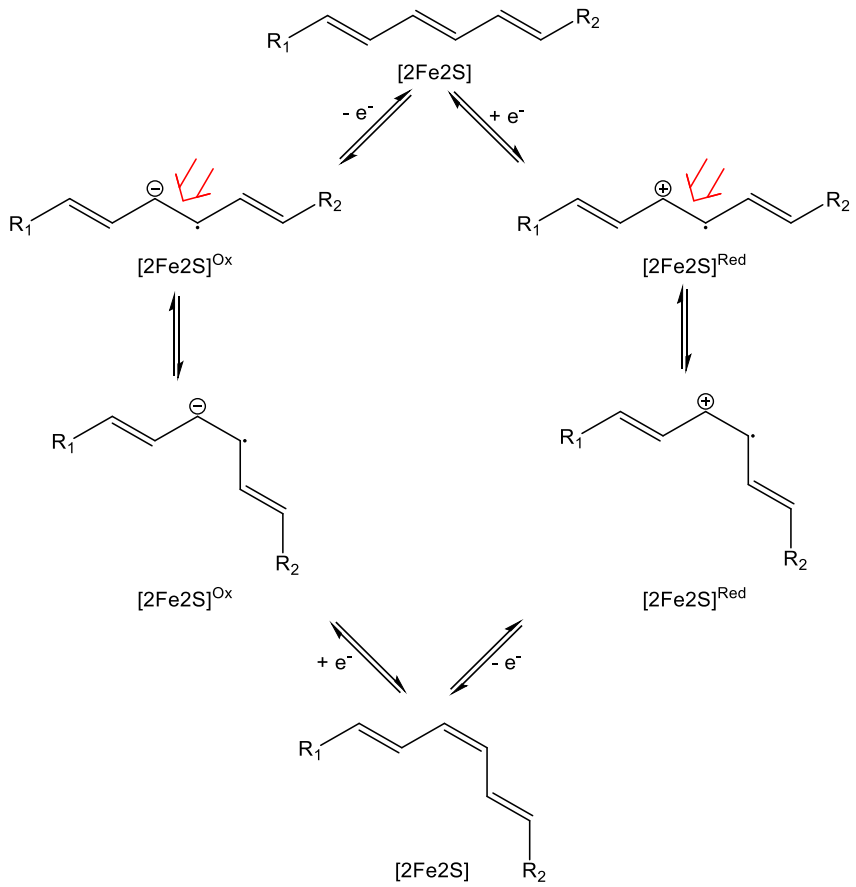


Figure 4.17 – Possible isomerisation mechanisms for the isomerisation of *β*-carotene by D27. Formation of either the substrate radical anion or radical cation would result in either an oxidised or reduced iron sulfur cluster. The resulting loss of conjugation in the substrate would allow free rotation of the newly formed σ bond. Reversal of the movement of the electron would reform the iron sulfur cluster and result in isomerised substrate. Mechanism involving a radical anion is more likely given the difficulties involved in abstraction of an electron from a conjugated system.

In order to investigate the presence of an [2Fe2S] cluster, GST-OsD27.1 was assayed in the presence of the 100 µM sodium dithionite, 100 µM sodium dithionite/pyridine, 100 µM 2-mercaptoethanol and 100 µM hydrogen peroxide. In the presence of each of these redox

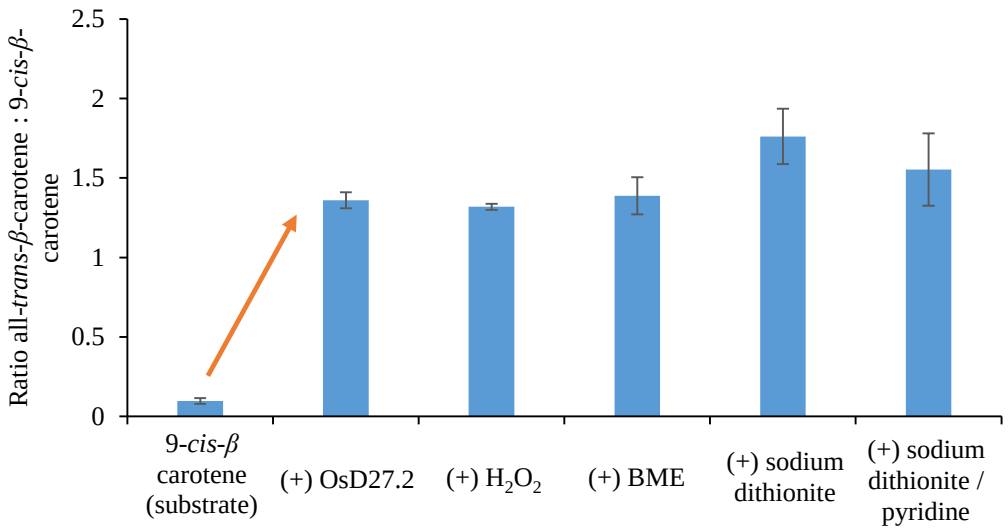


Figure 4.18 – Graph showing the ratio of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene following incubation of 9-*cis*-*β*-carotene with OsD27.2 in the presence of 100 µM hydrogen peroxide (H_2O_2), 100 µM 2-mercaptoethanol (BME), 100 µM sodium dithionite and 100 µM sodium dithionite and pyridine.

agents D27 was still active. D27 showed slightly higher activity in the presence of reducing agents than in the presence of hydrogen peroxide (Figure 4.18). However, the difference in the level of activity is not so significant as to constitute the presence of an iron sulfur cluster. Given the lack of variation in activity in the presence of oxidants and reductants, and given that D27 is active in aerobic conditions, if an [2Fe2S] cluster is present in D27, then it must be buried deep within the enzyme. This would render the cluster stable and protect it from denaturation by cytosolic components. This is consistent with the observation that D27 was still active in the presence of 100 µM of the metal chelators EDTA and 1, 10-phenanthroline. Commonly in proteins containing iron sulfur clusters the cluster is located on the exterior of the enzyme to allow electron transfer from the cluster to a second protein. Such is the case with Rieske dioxygenases, where a reductase is required to reduce a ferredoxin protein, which transfers an electron to an oxygenase, which performs catlaysis (Ferraro *et al*. 2005). In these cases, the solvent exposed iron-sulfur cluster of the ferredoxin is susceptible to oxidation, which can deactivate the iron sulfur protein (Orme-Johnson & Orme-Johnson 1978). It would be logical in the case of D27 for the cluster to be located on the interior of the enzyme, as this would allow the active site architecture of the enzyme to control the geometry of the isomer

produced. If the cluster was located on the exterior of the enzyme then no such control of the reaction would be possible, leading to a larger number of isomers being produced. As it is, although D27 is not regiospecific, it is predominantly selective for the 9-*cis* and all-*trans* isomers of *β*-carotene.

The UV absorption spectrum of D27 was recorded and found to be consistent with that recorded by Lin *et al*, with a distinct absorption band present at 422 nm and a small shoulder peak at 368 nm (Figure 4.19) (Lin *et al*. 2009). It is very difficult to discern the type of iron sulfur cluster present based on optical spectra. The absorption spectrum obtained for D27 is not consistent with plant type [2Fe2S] ferredoxins, which have additional absorption maxima at 280 nm, 330 nm, 420 nm and 460 nm (Ohme-Johnson 1973). Similarly other [2Fe2S] proteins have been shown to have absorption maxima at different positions (Table 4.3) (Palmer & Sands 1966, Tagawa & Arnon 1968, Fee *et al*. 1983). The shape of absorption spectra for D27 is itself, however, similar to that of NADH-quinone oxidoreductase from *Thermus thermophiles*, which has an absorption maxima at 410 nm for an $[4Fe4S]^{2+}$ cluster (Nakamaru-Ogiso *et al*. 2001). Calculation of the extinction coefficient at 422 nm for the D27 peak reveals a molar extinction coefficient of 6800 M^{-1} $cm^{-1,}$ which is consistent with [2Fe2S]

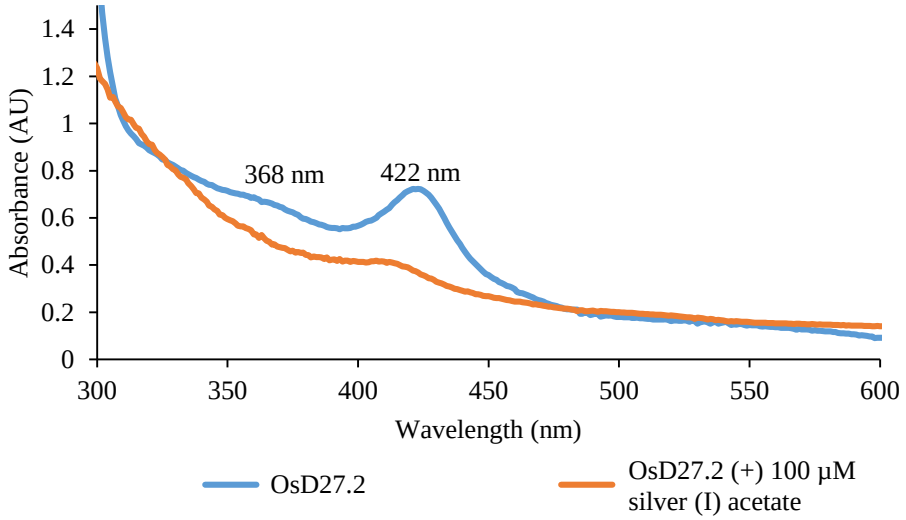


Figure 4.19 – UV spectra of OsD27.2 (blue) and OsD27.2 in the presence of 100 µM silver (I) acetate (orange). Upon the addition of silver acetate there is a reduction in the intensity of the peaks at 368 nm and 422 nm.

	Cluster Type	λ_{MAX} (nm)	Extinction coefficient (M^{-1} cm^{-1})	Reference
Plant type ferrodoxins (*Spinacia oleracea*)	[2Fe2S]	**280**, **330**, 420, 460	-	Palmer & Sands 1966
Higher Plant [2Fe2S]	[2Fe2S]	**325**, 420, 465	8000	Tagawa & Arnon 1968
T. thermophilus Rieske iron sulfur cluster	[2Fe2S]	380, **425**, 550	-	Fee *et al.* 1983
T. thermophilus NADH-quinone oxidoreductase	$[4Fe4S]^{2+}$	**410**	13300 – 15000	Nakamaru-Ogiso *et al.* 2001
O. sativa D27	?	368, **422**	6800	-

Table 4.3 – Comparison of λ_{MAX} values and extinction coefficients of selected different iron sulfur clusters from iron sulfur proteins. The dominant peak in the UV spectra for each protein is highlighted in bold.

(Ohme-Johnson 1973, Nakamaru-Ogiso *et al*. 2001). It appears likely, therefore, the iron sulfur cluster would be of the [2Fe2S] type.

Addition of 100 µM hydrogen peroxide, 100 µM 2-mercaptoethanol and 100 µM sodium dithionite did not affect the appearance of the peak at 422 nm. Lin *et al.* observed a reduction in the intensity of the peak at 422 nm in OsD27 protein in the presence of sodium dithionite (Lin *et al*. 2009). Addition of 100 µM silver (I) acetate, however, was found to cause a loss of intensity in the peak at 422 nm (Figure 4.19). Additionally, assays of OsD27 in the presence of 100 µM silver (I) acetate resulted in inhibition of enzyme activity. Interestingly, however, incubations in the presence of 100 µM lead (II) acetate did not result in a loss of activity. Silver (I) ions have been shown to destroy [4Fe4S] clusters in iron sulfur dehydratase enzymes in *E. coli* by displacing the iron from the cysteine ligands (Xu & Imlay 2012). Sodium dithionite similarly is known to quench iron sulfur clusters and demonstrates that the cofactor in D27 is redox active (Nakamaru-Ogiso *et al*. 2001). Taken together, the loss of a peak at 422 nm in the UV spectrum in the presence of silver (I) acetate and sodium dithionite is consistent with the presence of an iron sulfur cluster.

In conclusion, these observations provide some evidence of the presence of an iron sulfur cluster, possibly of the [2Fe2S] type (Figure 4.20). The oxidation state of the cluster is also unknown and due to the availability of only small amounts of enzyme, we were unable to perform further experiments, such as cyclic voltammetry. Further experimentation is required to ascertain the definitive presence of a cluster and the resting oxidation state of the cluster.

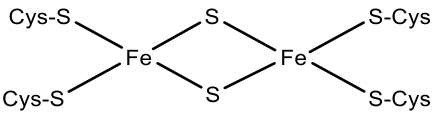


Figure 4.20 – Possible structure of the putative D27 active site [2Fe2S] iron sulfur cluster. The oxidation state of the cluster is unknown.

A BLAST search using the full length OsD27 sequence reveals a number of related unidentified proteins. Subsequent sequence alignment (Figure 4.21) of these proteins reveals large areas of conserved sequence between these proteins. This suggests that the structure of this region is essential for enzyme function. Within this highly conserved region it is possible to identify six conserved cysteine residues (Cys_{198}, Cys_{206}, Cys_{210}, Cys_{214}, Cys_{241} and Cys_{260} in *O. sativa*), which could potentially ligate an iron sulfur cluster.

4.7 Inhibition Assays of *O. sativa* Dwarf27

Sergeant *et al*. reported an *in planta* phenotypic effect of hydroxamic acids (compounds D2, D4, D5 and D6) on shoot branching in *Arabidopsis thaliana* (Sergeant *et al*. 2009). Shoot branching in plants is known to be mediated through the action of the phytohormone strigolactone (Gomez-Roldan *et al*. 2008, Umehara *et al*. 2008). The complete biosynthetic pathway of strigolactone is unknown. However, the first three enzymes in the biosynthetic pathway are known: Dwarf27, carotenoid cleavage dioxygenase 7 and carotenoid cleavage dioxygenases 8. Although hydroxamic acids are designed to target the carotenoid cleavage dioxygenases, it was decided to test the hydroxamic acids *in vitro* against D27 in order to investigate whether the hydroxamic acids affect D27 and to discern the selectivity of the hydroxamic acids towards the different enzymes of the strigolactone biosynthesis pathway. To this end, OsD27.2 was screened against the hydroxamic acids (structures shown in tables

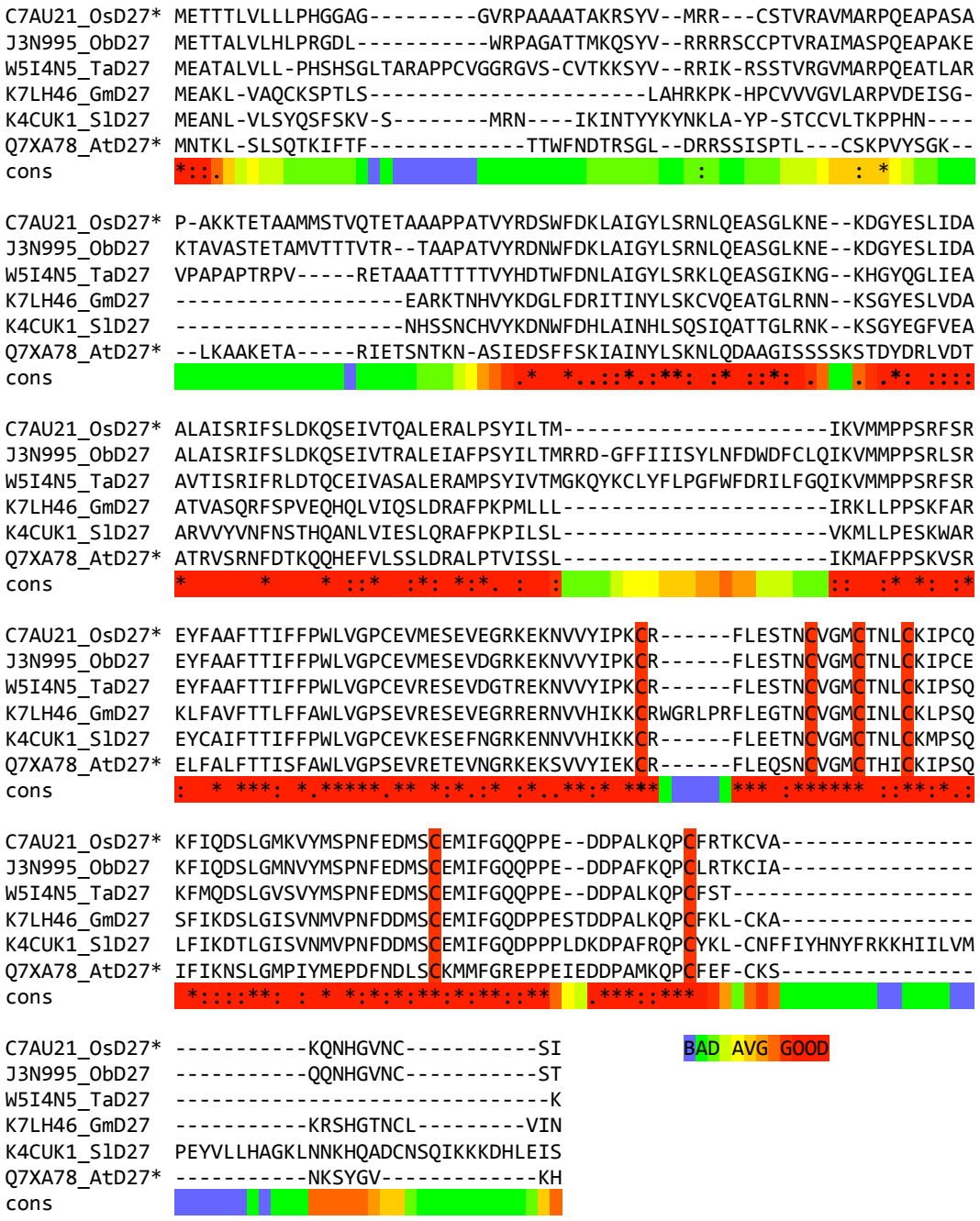


Figure 4.21 – T-Coffee sequence alignment of OsD27 against selected random hits from a BLAST search. Conserved cysteine residues are highlighted in red. Ob – *Oryza brachyantha* (wild rice); Ta – *Triticum aestivum* (wheat); Gm – *Glycine max* (soybean) *functionally characterised enzyme.

2.1-2.5 and 3.3-3.7). The hydroxamic acids D12H and D13H (synthesised by Paul Sainsbury), abamine(SG) (provided by Syngenta) and nordihydroguaiaretic acid (Sigma) were also assayed at 100 µM (Figure 4.22). All inhibitors were assayed at a final concentration of 100 µM in the presence of 50 µg total protein containing D27 and 20 µM 9-*cis*-*β*-carotene. All incubations were performed for 20 minutes at 25° C in the dark.

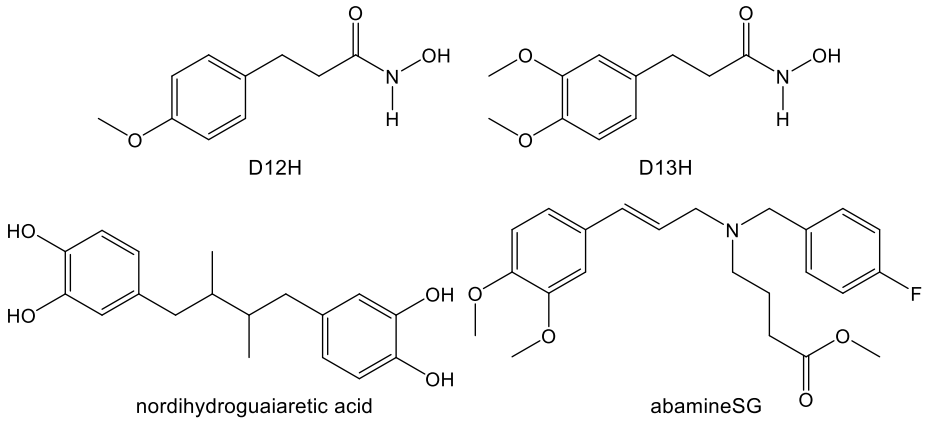


Figure 4.22 – Structures of additional inhibitors screened against OsD27.2– D12H, D13H, nordihydroguaiaretic acid (NDGA) and abamineSG.

None of the hydroxamic acids tested showed complete inhibition of OsD27.2 at 100 µM. However, some hydroxamic acids, as well as NDGA, did show some inhibition (Figure 4.23, Tables 4.4 and 4.5). In addition, no inhibition of D27 activity was reported for abamine or abamineSG at 100 µM. These data are unsurprising, given that OsD27 is suspected to be an iron sulfur protein. Since hydroxamic acid activity would likely result from hydroxamic acid chelation to the active site iron, thus blocking substrate access, it is likely that the hydroxamic acids are not chelating to any active site iron in the CCDs. This could be due to the architecture of the iron sulfur cluster. If the iron is coordinated by two bridging sulfur atoms and 4 cysteine residues (Figure 4.19), then this may not provide any free coordination sites onto which the hydroxamic acid can coordinate, thus preventing inhibition of the enzyme.

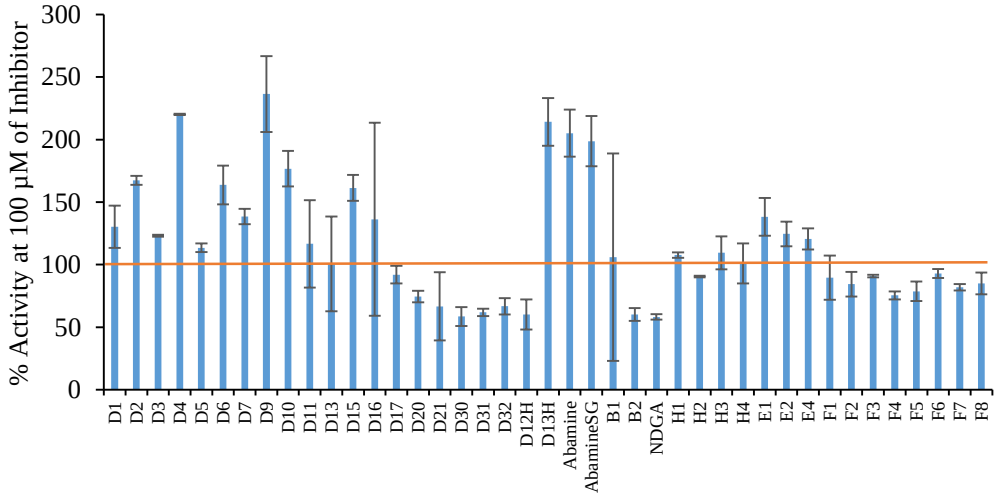


Figure 4.23 – Percentage activity of OsD27.2 isomerisation of 9-*cis*-*β*-carotene in the presence of hydroxamic acid inhibitors.

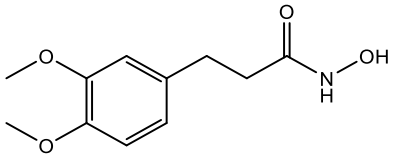
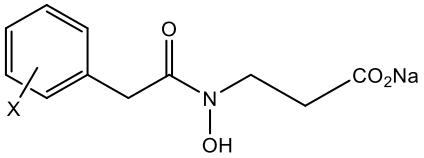
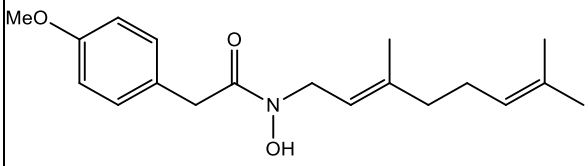
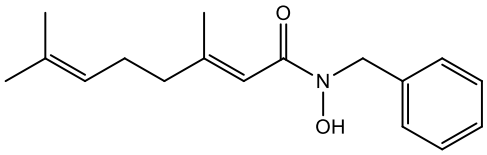
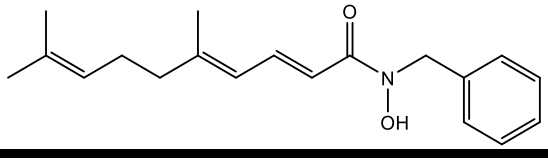
Inhibitor	Structure	Substituent X	OsD27 Inhibition at 100 µM
D12H		---	40.0
D20		4-Methoxy	25.5
D21		Napthyl	33.3
D30		---	41.4
D31		---	38.1
D32		---	33.3

Table 4.4 – Inhibition data for hydroxamic acid compounds D12H, D20-D21 and D30-D31 versus OsD27.

It has been observed that hydroxamic acids bind with a greater affinity to iron (III) (Sergeant *et al*. 2013). Since it is unknown what oxidation state the iron is in, it is possible that the hydroxamic acids are binding to D27, but with an affinity too low to induce inhibition. Similarly it may well be that the active site architecture precludes the hydroxamic acids from binding.

Compounds D30, D31 and D32 all showed approximately 33% to 42% inhibition of OsD27.2 at 100 µM. These compounds are terpenoid-like, mimicking the terpenoid structure of the carotenoid substrates. It is possible that this architecture allows D30, D31 and D32 to bind to the active site without necessarily chelating the active site iron. Similarly, B2 and H1 are a longer compounds compared to other members of the hydroxamic acid inhibitors.

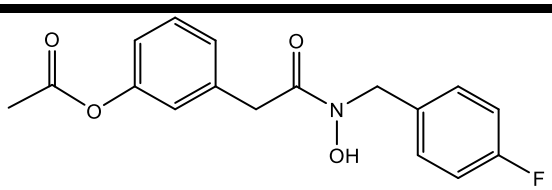
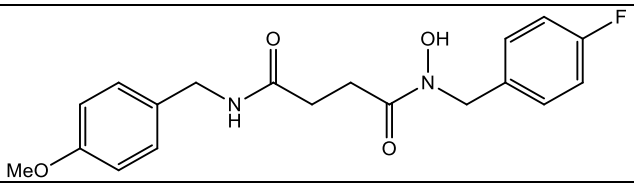
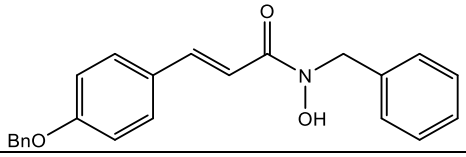
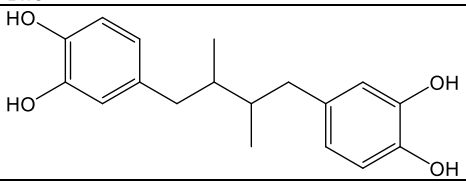
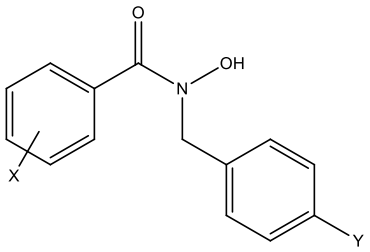
Inhibitor	Structure	Substituent X	Substituent Y	OsD27 Inhibition at 100 µM
D17		---	---	8.0
B2		---	---	39.7
H2		---	---	9.4
NDGA		---	---	41.8
F1		4-Methoxy	H	10.4
F2		4-Methoxy	F	15.5
F3		3,4-Dimethoxy	H	9.1
F4		3,4-Dimethoxy	F	24.5
F5		3-Chloro	H	21.3
F6		3-Amino	H	7.0
F7		3-Bromo	H	18.0
F8		3-Chloro-4-Methoxy	H	15.0

Table 4.5 – Inhibition data for hydroxamic acid compounds F1-F8 versus OsD27. Bn represents a benzyl group (CH_2Ph).

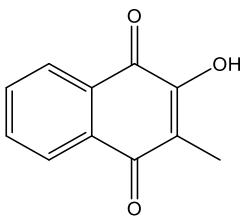
Likewise, they could be inhibiting D27 through non-specific binding in the active site. Both D20 and D21 showed some inhibition of OsD27.2, 25% and 33% respectively. Both D20 and D21 have carboxylate groups, which are likely to be deprotonated at reaction pH (pH 6.4). These carboxylate groups could be interacting with residues in the active site or to the iron sulfur cluster itself, blocking the active site and resulting in inhibition. D17 also exhibited a

small level of inhibition, although based on the inhibition observed with other members of the hydroxamic acid inhibitors, it is not clear why D17 would show inhibition. As such, it appears that larger hydroxamic acids are better inhibitors of D27 through non-specific binding to the D27 active site.

However, inhibition was also observed for all members of the 'F' series compounds and D12H. Compounds F1 – F8 and D12H are small compounds which showed between 7% and 24% (for F1 – F8) and 40% (for D12H) inhibition of D27, in contradiction to the observation that larger hydroxamic acids showed inhibition. However, it is possible that an inverse effect is occurring, whereby the small nature of these compounds allows them access to a small binding pocket within the active site, blocking substrate access and inhibiting the enzyme. It is interesting that whilst D12H showed inhibition, there was no inhibition by D12, D13 and D13H, which are structurally related compounds. Likewise, it is unlikely that inhibition is the result of chelation by the hydroxamic moiety, and is more likely due to non-specific interactions within the active site of the enzyme.

The compound NDGA showed 42% inhibition against OsD27.2. NDGA is similar to B2, D30, D31 and D32 in that it is a long molecule. However, NDGA also contains two catechols. These catechols may bind to the iron sulfur cluster, resulting in enzyme inhibition.

A number of quinone based compounds have been shown to have inhibitory effects on electron transport chains, which use iron sulfur clusters to mediate electron transfer. These quinones inhibit these proteins by inhibiting electron transfer (Matsuura *et al*. 1982). Based on this, 2-hydroxy-3-methyl-1,4-napthoquinone (Sigma) (Figure 4.24) was purchased and



2-hydroxy-3-methyl-1,4-napthoquinone

Figure 4.24 – Structure of 2-hydroxy-3-methyl-1,4-napthoquinone.

assayed against OsD27.2 at 100 µM. 2-hydroxy-3-methyl-1,4-napthoquinone is a member of a family of quinones which have been shown to inhibit electron transfer in electron transport chains. However, no inhibition was observed. This absence of inhibition does not indicate an absence of an iron sulfur cluster, as the lack of inhibition may well be due to inability of the 3-alkyl-2-hydroxy-1,4-napthoquinone to enter the active site of the enzyme.

Only one compound was found to fully inhibit OsD27.2 isomerisation of 9-*cis*-*β*-carotene to all-*trans*-*β*-carotene: silver (I) acetate. Inhibition by silver (I) acetate likely arises due to displacement of iron in the active site iron sulfur cluster by silver. Interestingly lead (II) acetate did not inhibit the activity of OsD27.2. The reasons for this are unknown, but could be due to the larger size of the lead ion.

4.8 Conclusions

On the basis of the data obtained and observations in the literature, D27 is a putative iron sulfur enzyme containing an [2Fe2S] iron sulfur cluster. D27 is sparingly soluble, indicating that it is a membrane associated protein, which is to be expected given the hydrophobic nature of the carotenoid substrates which are found localised within membranes. Given that D27 is active independent of any cofactors or other cytosolic components this indicates D27 is not dependent upon electron transport chain proteins like CRTISO or redox cofactors such as NAD(H) or FAD(H). However, the data obtained do not allow for a definitive assignment of the presence of such an iron sulfur cluster, and additional experiments will be required to demonstrate the presence of an iron sulfur cluster. D27 appears to exhibit substrate inhibition by 9-*cis*-*β*-carotene, formed from the isomerisation of all-*trans*-*β*-carotene.

The isomerisation of 9-*cis*-*β*-carotene to all-*trans*-*β*-carotene by D27 is only partially inhibited by the action of the hydroxamic acid inhibitors *in vitro*, and it is not possible to discern any definitive structure-activity relationships from those compounds that do show some inhibition. Moreover, no inhibition was seen for compounds D2, D4, D5 and D6, which

exhibit a shoot branching phenotype *in planta*. These results suggest that the effects on shoot branching must be due to the inhibition of another target, most likely CCD7 or CCD8.

4.9 Future Work

D27 is a remarkable enzyme which catalyses a highly specific isomerisation reaction with remarkable precision via an unusual enzymatic mechanism. Elucidation of this mechanism would be of wider interest and, assuming the presence of an iron sulfur cluster, would demonstrate a novel function of an iron sulfur protein. A key experiment for determining the presence of such an iron sulfur cluster would be cyclic voltammetry, which would reveal whether a redox process is occurring during the isomerisation of β-carotene by D27. Additionally, mass spectrometry of the native protein would also aid in this endeavour. Upon the treatment of an iron sulfur protein with mild acid, the iron sulfur cluster should break down due to competition for the cysteine ligands from protons. This breakdown of the iron sulfur cluster could be detected by a change in mass. Site directed mutagenesis of conserved key cysteine residues could also be used to identify whether the cysteine residues are required for catalysis. Mossbauer spectroscopy and EPR spectroscopy could also be used to identify the presence of an iron sulfur and help to identify the type of cluster present.

It would be enlightening to examine the preference of D27 for different carotenoids. As has been discussed previously in Chapter 3, the biosynthesis of the hormone abscisic acid requires isomerisation of all-*trans*-neoxanthin to 9'-*cis*-neoxanthin. As of yet, the enzyme or activity which performs this function has yet to be identified. A BLAST search on the UniProt database reveals that there are four orthologs of D27 present in *A. thaliana* (Entry numbers: Q7XA78, Q9SGU7, F4HZB1 and Q8W4C4) with 53%, 38% 49% and 32% similarity to OsD27 respectively. It is entirely probable that one of these homologs could have different substrate specificity. Developing a D27 specific inhibitor would aid in this, and would allow additional phenotypes mediated by these types of enzymes to be identified.

Chapter Five: Expression and Inhibition of *Arabidopsis thaliana* Carotenoid Cleavage Dioxygenase 7 (CCD7)

5.1 Introduction

Carotenoid cleavage dioxygenases 7 (CCD7) is the second enzyme on the committed biosynthesis pathway of strigolactone, cleaving 9-*cis*-*β*-carotene to 9-*cis*-*β*-apo-10'-carotenal (Alder *et al*. 2012) (Figure 5.1). CCD7 will also cleave the all-*trans* isomer of *β*-carotene, albeit ten fold less efficiently. The 9-*cis* isomer is reported to be the preferred natural substrate for the enzyme (Alder *et al*. 2012). The CCD7 cleavage product, 9-*cis*-*β*-apo-10'-carotenal, is then converted by CCD8 to carlactone and subsequently by a number of other enzymes to strigolactone. CCD7 has also been shown to cleave the 9-*cis* isomers of zeaxanthin and lutein (Bruno *et al*. 2014). Other than this, there is little other biochemical information known about CCD7.

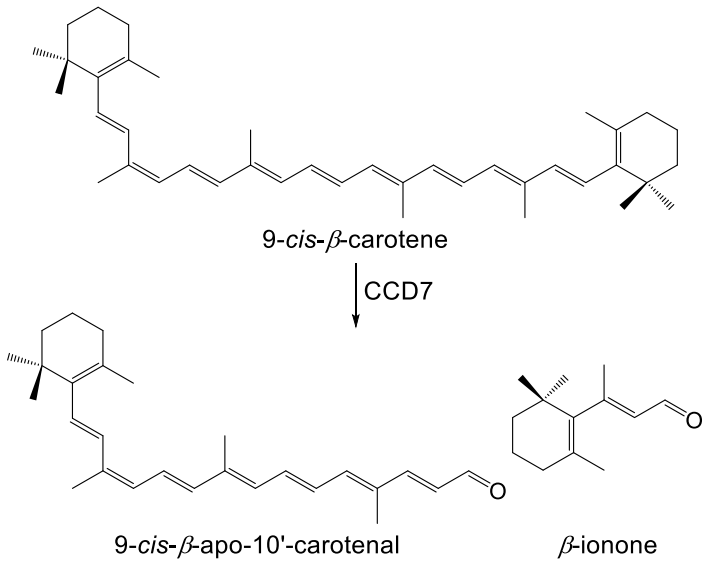


Figure 5.1 – Schematic of the cleavage reaction catalysed by CCD7.

The work detailed within this chapter describes the expression of *Arabidopsis thaliana* CCD7 and assays of *A. thaliana* CCD7 against hydroxamic acid inhibitors to determine the biochemical basis of observed *in planta* phenotypes in the presence of

hydroxamic acids. Work on the development of the AtCCD7 assays was completed in conjunction with Erasmus project student Flora Descombes.

5.2 Expression of *A. thaliana* CCD7

Two constructs, pGEX-AtCCD7.1 (full length CCD7 gene) and pGEX-AtCCD7.2 (truncated gene lacking the first 93 base pairs corresponding to a chloroplast transit peptide), both with a GST fusion tag, were obtained from Dr Andrew Thompson (Cranfield University) (CCD7 accession number: NM_130064. The sequence of construct used differs slightly from that of the wild type and the sequence of CCD7 used is shown in Appendix 10.5). Both plasmids were transformed into competent pRosetta BL21 *E. coli* and the GST-AtCCD7.1 and GST-AtCCD7.2 fusion proteins were produced following induction with 100 µM IPTG at A_{600} = 0.6 AU, and grown overnight at 20° C. However, following sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of the cell free extract and fractions following glutathione-S-transferase (GST) tagged purification, no protein was visible in either the total protein or elution fractions. Varying the temperature of protein induction and the concentration of IPTG present did not result in the appearance of a band of molecular weight which would correspond to either AtCCD7.1 or AtCCD7.2 (95.9 and 92.6 kDa respectively). At this point, given the tendency of signal peptide sequences to cause aggregation and interfere with protein expression, only work on AtCCD7.2 was continued (Singh *et al*. 2013).

The pBB541 plasmid (see Chapter 4.2) was used to transform competent pRosetta BL21 *E. coli*. The resulting transformant was then made chemically competent via the method of Inoue and the pGEX-ATCC7.2 plasmid was subsequently used to transform the pRosetta pBB541 BL21 *E. coli* (Inoue *et al*. 1990). Upon co-expression of GroES and GroEL with AtCCD7.2 a band of the correct molecular weight, 92.6 kDa, was visible in the SDS-PAGE gel (Figure 5.2).

The glutathione-S-transferase tagged AtCCD7.2 fusion protein was purified by glutathione affinity chromatography and the partially purified enzyme was used for assays

(Figure 5.2). Due to the low expression of AtCCD7.2 further purification was not attempted and the partially purified protein was used as described following desalting from purification elution buffer.

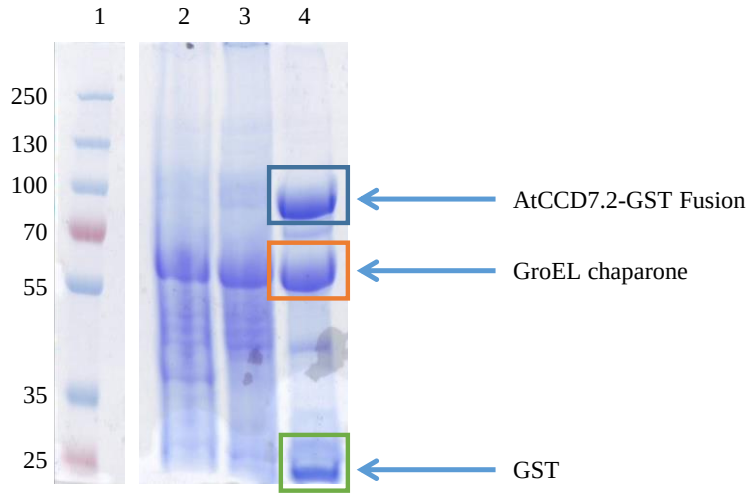


Figure 5.2 – 8% SDS-PAGE gel showing the overproduction of *A. thaliana* CCD7.2 from pGEX-AtCCD7.1 pRosetta pBB541 BL21 *E. coli*. 1 – Ladder; 2 – Total protein; 3 – Total soluble protein; 4 – Elution fraction following concentration of protein. Molecular weights of ladder are in kDa. Molecular weight of GST tagged AtCCD7.2: 95.9 kDa (calculated from sequence). Mass spectrometry would be required to definitively prove the presence of the AtCCD7 protein.

5.3 Assays of *A. thaliana* CCD7

Initially, assays were attempted by Flora Descombes with commercially available all-*trans*-*β*-carotene and cell free extract containing overproduced AtCCD7.2. In these assays all-*trans*-*β*-carotene was solubilised in 4% (v/v) Triton X-100 in ethanol and added to a final concentration of 40 µM. The reaction was allowed to proceed for two hours in the light. Following extraction into ethyl acetate the products were analysed via C_{30} reversed phase HPLC at room temperature (as described in Chapter 4). These initial assays showed no turnover of the *β*-carotene to the apocarotenoid product (Figure 5.3). Given the impurity of 9-*cis*-*β*-carotene in commercial all-*trans*-*β*-carotene (as discussed in Section 4.3), and the fact that reactions were incubated for two hours in the light to allow light mediated isomerisation of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene, it was expected that turnover would have been observed despite the ten fold slower reaction of CCD7 with all-*trans*-*β*-carotene over 9-*cis*-*β*-carotene.

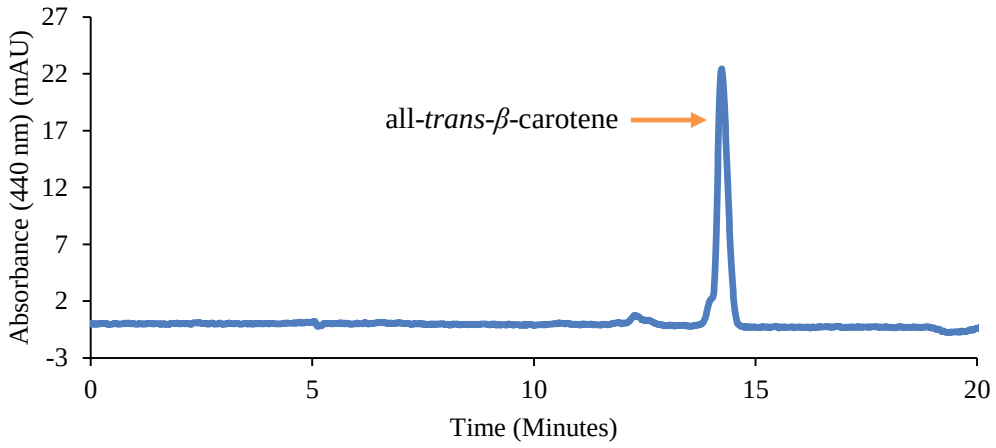


Figure 5.3 – HPLC chromatogram at 440 nm following incubation of commercial all-*trans*-*β*-carotene with AtCCD7.2. No new peaks are visible at 440 nm.

For assays with 9'-*cis*-epoxycarotenoid cleavage dioxygenases (NCED), using the 9'-*cis*-neoxanthin substrate, the products are extracted into ethyl acetate (Chapter 3). However, both *β*-carotene and the subsequent aldehyde and ketone products produced from the cleavage reaction are much less polar, and may not extract well into ethyl acetate. Assuming the apparent lack of turnover of AtCCD7.2 was due to lack of uptake of the products into ethyl acetate, the extraction medium was changed to dichloromethane, a significantly more non-polar solvent. On repeating the experiment with dichloromethane extraction, some possible product peaks were observed in the HPLC chromatogram. A peak at 3.9 minutes was observed, which absorbed at 300 nm, corresponding to the *β*-ionone product, matching the retention time of authentic *β*-ionone standards (λ_{MAX} of *β*-ionone: 299 nm) (Figure 5.4). Additionally, a new small peak was observed at 12.5 minutes, absorbing at 440 nm. However, it was unknown whether this peak corresponded to either 9-*cis*-*β*-apo-10'-carotenal or all-*trans*-*β*-apo-10'-carotenal, since authentic standards are not available, or whether it was another isomer of *β*-carotene.

It was decided to chemically generate 9-*cis*-*β*-carotene from all-*trans*-*β*-carotene as it was reasoned that the apparent lack of turnover may be due in part to a low concentration of substrate (the 9-*cis* isomer) and the low turnover rate of AtCCD7.2 towards the all-*trans* isomer. Additionally at this point, 10 µL of partially purified AtCCD7.2 was used in assays as

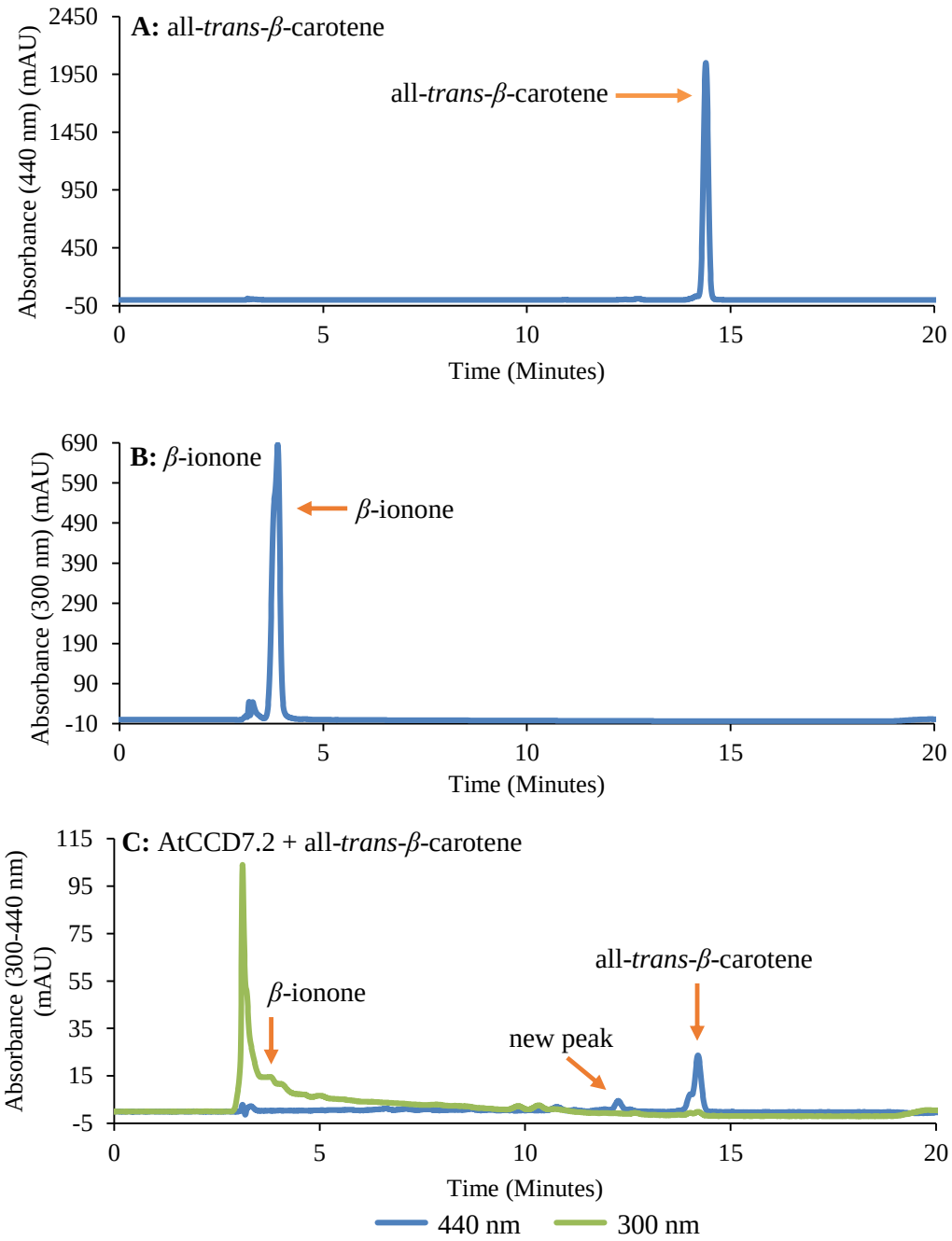


Figure 5.4 – HPLC chromatogram showing all-*trans*-*β*-carotene (**A**) and *β*-ionone (**B**) controls and following incubation of all-*trans*-*β*-carotene with AtCC7.2 (**C**). Upon incubation of all-*trans*-*β*-carotene with AtCCD7.2 new peaks at 3.9 minutes (corresponding to the *β*-ionone) and 12.5 minutes are visible.

opposed to cell free extract (assays contained approximately 125 µg of protein, as determined by Bradford assay, following purification (Bradford 1976)). The 9-*cis* isomer of *β*-carotene was generated via iodine catalysed isomerisation in the presence of ultraviolet light from all-*trans*-*β*-carotene (Figure 5.5). The isomerisation reaction produces a mixture of different

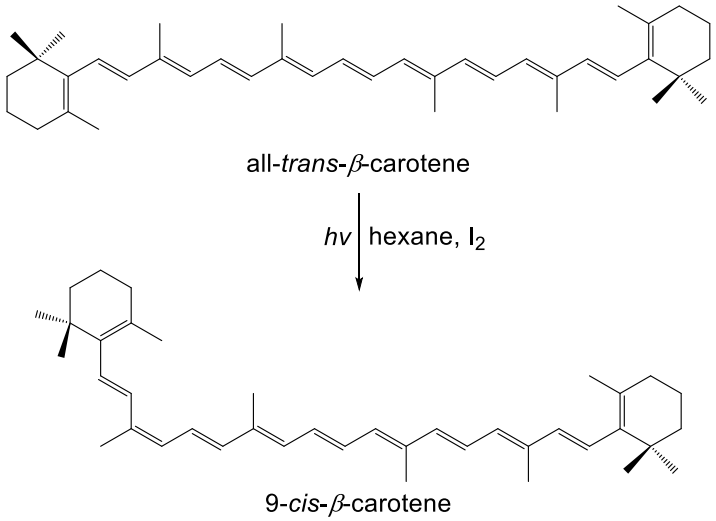


Figure 5.5 – Iodine catalysed isomerisation of all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene (and other isomers not shown).

geometrical isomers which can be purified via reversed phase HPLC on a C_{30} column, at 0° C (Figure 5.6). Following collection of the appropriate peak (at 23.3 minutes, as described in Section 4.3) the solvent was removed under reduced pressure and the 9-*cis*-*β*-carotene was resuspended in ethanol. Using a known weight of commercially available all-*trans*-*β*-carotene, the extinction coefficient of *β*-carotene at 475 nm was calculated as 6.55 x 10^{-4} M^{-1} cm^{-1}. This extinction coefficient was used to calculate the concentration of 9-*cis*-*β*-carotene purified (Beer 1852). It should be noted that the 9-*cis*-*β*-carotene used was not completely geometrically pure due to isomerisation which occurs as discussed in Chapter 4. Other isomers present include the all-*trans* and 15-*cis* isomers.

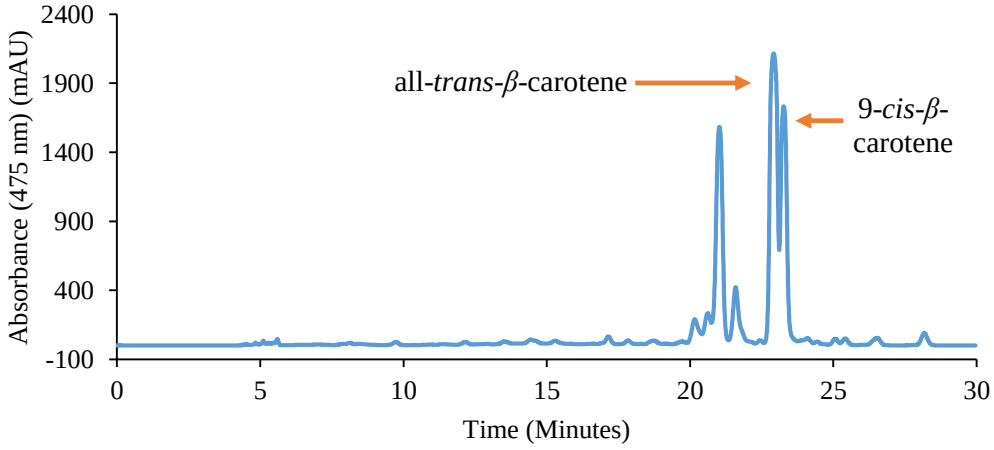


Figure 5.6 – HPLC chromatogram at 475 nm of the purification of 9-*cis*-*β*-carotene from a geometrical mix of *β*-carotene isomers. All-*trans*-*β*-carotene is still the major isomer present.

After incubation of AtCCD7.2 in the presence of 40 µM 9-*cis*-*β*-carotene, a new peak was observed in the HPLC chromatogram at approximately 10.5 minutes at 440 nm (Figure 5.7). Control experiments in the presence of boiled AtCCD7.2 failed to produce a peak (Figure 5.7). Additionally the appearance of the peak proceeded in a time dependent fashion which correlated with the disappearance of the 9-*cis*-*β*-carotene (Figure 5.8) and HRMS gave the correct molecular formula for *β*-apo-10'-carotenal (Figure 5.8). However, it was unknown whether this peak belongs to the *cis* or *trans* isomer. Analysis of the UV spectrum of this peak showed a λ_{MAX} of 444 nm, in agreement with a value of 447 nm reported by Alder *et al*. for 9-*cis*-*β*-apo-10'-carotenal (Alder *et al*. 2012) (Figure 5.8). As a result of this, it was concluded that the presence of the peak at 10.5 minutes was due to the action of AtCCD7.2 and belonged to 9-*cis*-*β*-apo-10'-carotenal.

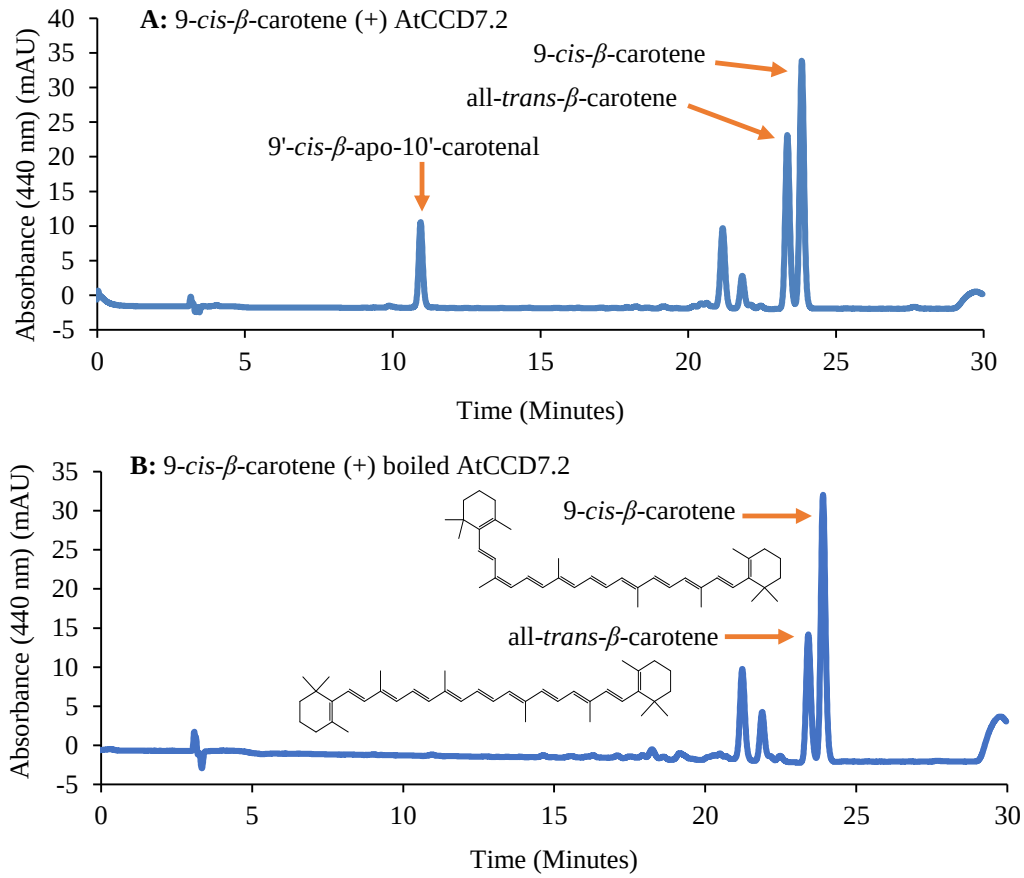


Figure 5.7 – **A:** HPLC chromatogram at 440 nm following incubation of 9-*cis*-*β*-carotene with AtCCD7.2. A new peak corresponding to 9-*cis*-*β*-apo-10'-carotenal is visible at 10.5 minutes. **B:** Control assay in the presence of boiled AtCCD7.2. New peaks at 20.2m and 21.8m are proposed to arise due to non-specific photochemical isomerisation.

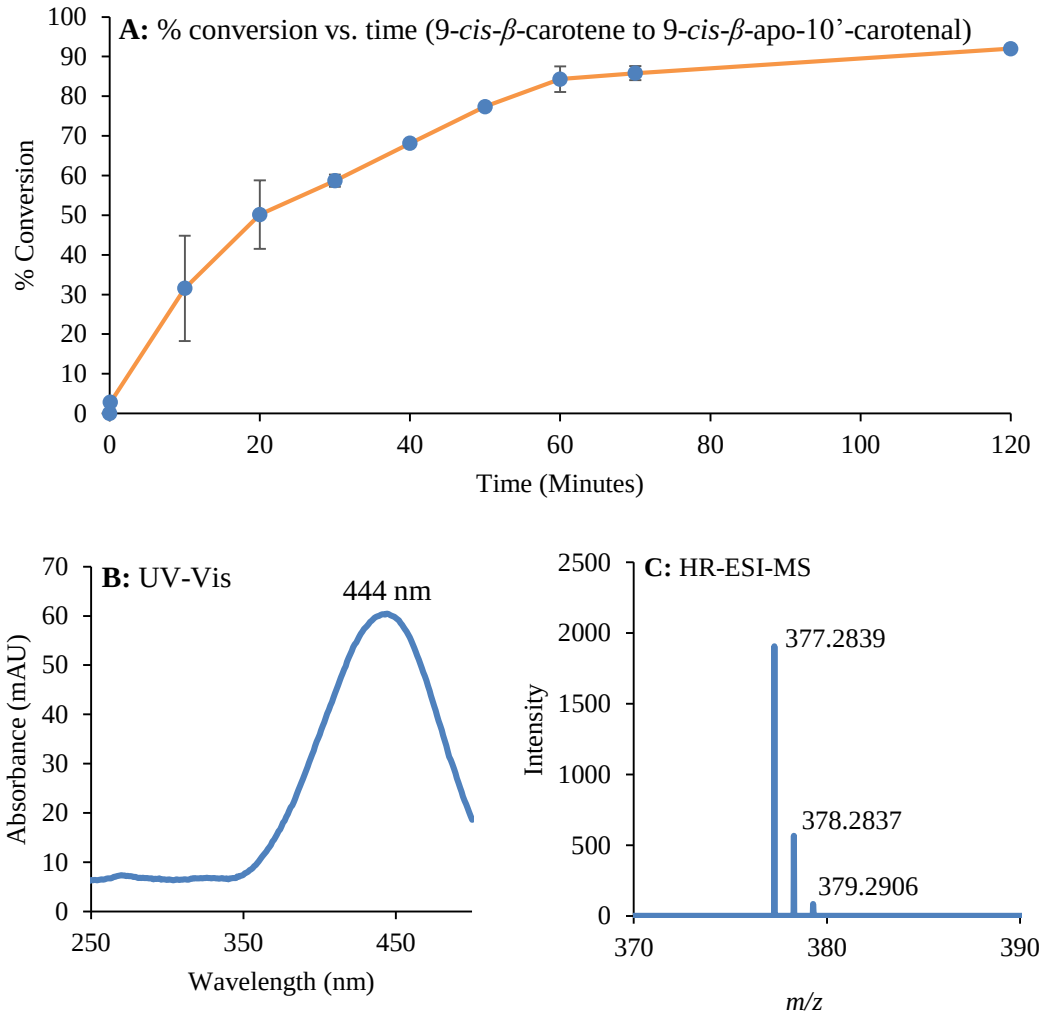


Figure 5.8 – Percentage conversion of 9-*cis*-*β*-carotene to 9-*cis*-*β*-apo-10'-carotenal versus time (**A**), UV-Vis (**B**) and ESI-MS (**C**) analysis of the peak at 10.5 minutes from Figure 5.8, confirming the identity as 9-*cis*-*β*-apo-10'-carotenal. ESI-MS: *m/z* calculated for $C_{27}H_{36}O+H^+$ 377.2839, found 377.2840.

The yield of the peak produced was rather low (≈30%), considering the incubation had been at 25° C for one hour, suggesting either inefficiency of the enzyme or a low substrate concentration. Other factors considered for the apparent low turnover of the enzyme were inhibition of the enzyme by other isomers of *β*-carotene or a low concentration of substrate. Using authentic standards of *β*-ionone it was known that the *β*- ionone product eluted at 3.9 minutes, although it co-eluted with other extraction products which also absorb at 300 nm. Additionally, *β*-ionone is a volatile compound, and it is likely that during evaporation of solvent during the extraction process a proportion of the *β*-ionone is lost, resulting in a decreased signal on the HPLC. It was noted, however, that the magnitude of the peak at 3.9

minutes was larger in the presence of AtCCD7.2 than in the absence of AtCCD7.2, strongly suggesting that the *β*-ionone was indeed produced as expected (Figure 5.9).

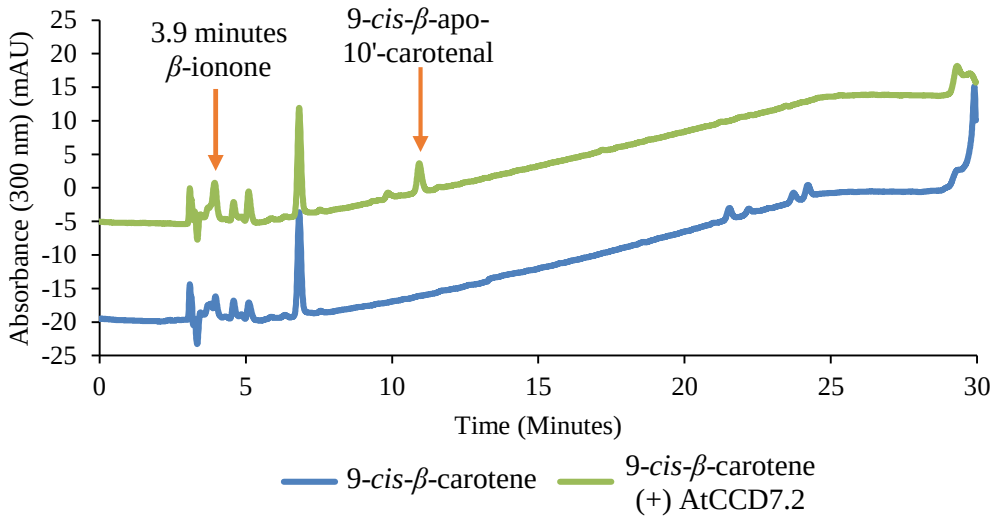


Figure 5.9 – Comparison of HPLC chromatogram at 300 nm of 9-*cis*-*β*-carotene (blue) and 9-*cis*-*β*-carotene plus AtCCD7.2 (orange). The peak at 3.9 minutes is greater in the presence of AtCCD7.2, likely due to production of *β*-ionone.

It was further observed that upon addition of either ethyl acetate or dichloromethane to the assay mix there was precipitation of protein, as would be expected. Following centrifugation to separate the organic ethyl acetate from the aqueous reaction mixture the protein precipitate had formed a layer between the organic and aqueous layers. The precipitate layer was coloured orange, suggesting the *β*-carotene and possibly the products were in the precipitate. It was reasoned that the *β*-carotene and hydrophobic products were co-precipitating with the protein upon addition of the organic solvent due to the hydrophobic environment created by the precipitated protein. This is unsurprising, given the hydrophobic nature of the substrate and products, and hence energetically it is more favourable for the hydrophobic components to precipitate with the protein. As a result it is likely that little of the substrate and products were being extracted into the organic solvents, the outcome being the lack of apparent turnover of the enzyme.

In order to overcome poor extraction of products, the ionic strength of the assay solution was increased following completion of the assay prior to extraction into organic solvent. This was achieved through the use of a solution of 4 M pyridinium acetate at pH 5.0,

produced by adding pyridine to acetic acid and water (Mihalyi 2014). When added to the assay mixture following completion of the assay, the pyridinium acetate increases the ionic strength of the solution from effectively zero to 4.01 M. This increased ionic strength not only results in denaturation of the proteins but also allows the charged protein molecules to stay in solution, preventing co-precipitation of the *β*-carotene. Subsequent addition of *n*-butanol results in extraction of the enzyme substrates and products. Pyridinium acetate could be substituted with 4 M sodium chloride solution pH 7.0 with the same effect. Due to the high salt concentrations, the *n*-butanol extract must be first desalted via C_{18} solid phase extraction cartridges. The substrate and products are then eluted into tertbutylmethylether, which is removed under reduced pressure. The residue can then be resuspended in methanol and analysed via reversed phase HPLC on a C_{30} column.

Following assays with 9-*cis*-*β*-carotene and extraction in pyridinium acetate, a peak corresponding to the 9-*cis*-*β*-apo-10'-carotenal was consistently detected at 10.5 minutes. It was found that enzyme incubations at 25° C with shaking at 180 rpm in the dark resulted in good activity of the enzyme.

5.4 Kinetic Characterisation of *A. thaliana* CCD7

Experiments were performed in order to determine the Michaelis constant (K_M) of AtCCD7.2. Using the stopped HPLC method, the K_M was calculated as 8.7 µM for 9-*cis*-*β*-carotene at pH 7.8 at 25° C (Figure 5.10). This compares moderately well with the value of 15.2 µM reported by Schwartz *et al*. for *β*-carotene (likely a geometrically impure mix, with predominately the all-*trans* isomer) but differs slightly from the value of 41.1 µM reported by Bruno *et al*. for 9-*cis*-*β*-carotene (Schwartz *et al*. 2004 & Bruno *et al*. 2014). Differences between the three reported values are likely due to the fact that the K_M must be determined through a stopped HPLC assay, in addition to variations in pH and temperature.

During the time course experiment of AtCCD7.2 it was noticed that the amount of all-*trans*-*β*-carotene and 15-*cis*-*β*-carotene decreased versus time, albeit at a slower rate than

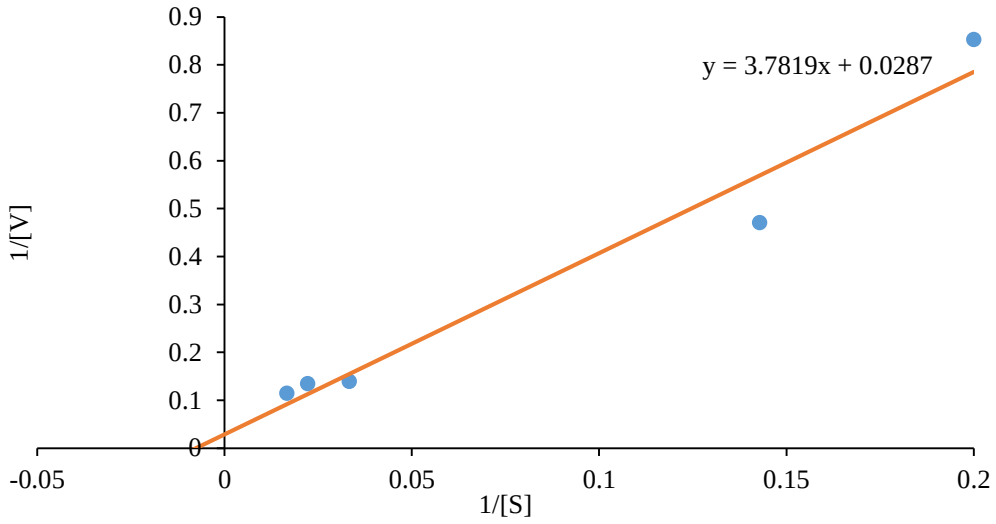


Figure 5.10 – Lineweaver-Burk plot for the determination of the K_M of AtCCD7.2 for 9-*cis*-*β*-carotene at pH 7.8 at 25° C.

the 9-*cis* isomer (Figure 5.11 and 5.12). In addition to this, a new small peak was noticed at a retention time of 9.8 minutes, with a λ_{MAX} of 457 nm (Figure 5.13). The reported λ_{MAX} for all-*trans*-*β*-apo-10'-carotenal is 452 nm and the all-*trans* isomer is known to elute before the 9-*cis* isomer (Alder *et al*. 2012). When the 9-*cis*-*β*-apo-10'-carotenal was purified (and shown to be pure via immediate reinjection to HPLC) and then exposed to heat (40° C) and UV light, the peak at 9.8 minutes reappeared following HPLC analysis, demonstrating the 9.8 minute

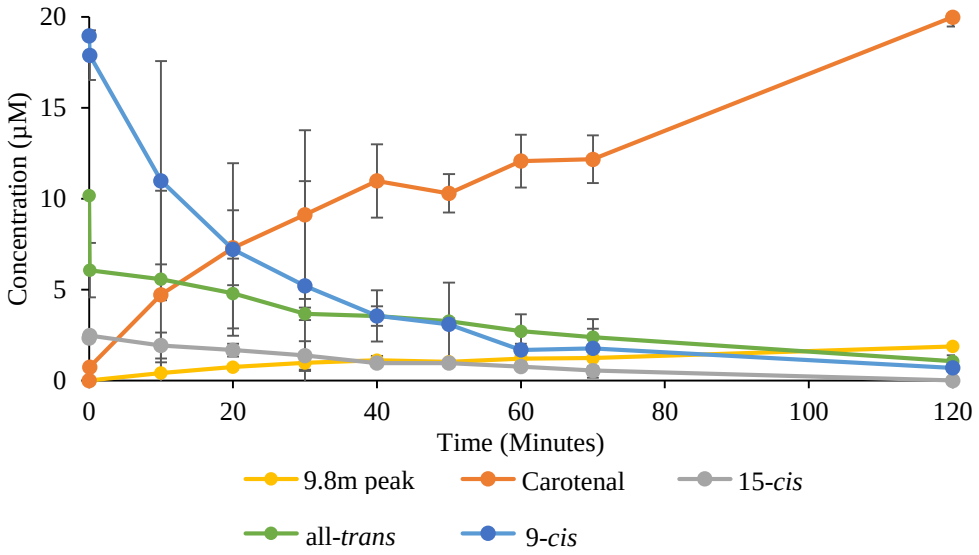


Figure 5.11 – Graph showing the change in concentration of all-*trans*-*β*-carotene (all-*trans*, green), 9-*cis*-*β*-carotene (9-*cis*, blue), 15-*cis*-*β*-carotene (15-*cis*, grey), 9-*cis*-*β*-apo-10'-carotenal (carotenal, orange) and the peak at 9.8 minutes (9.8m peak, yellow) (believed to belong to all-*trans*-*β*-apo-10'-carotenal).

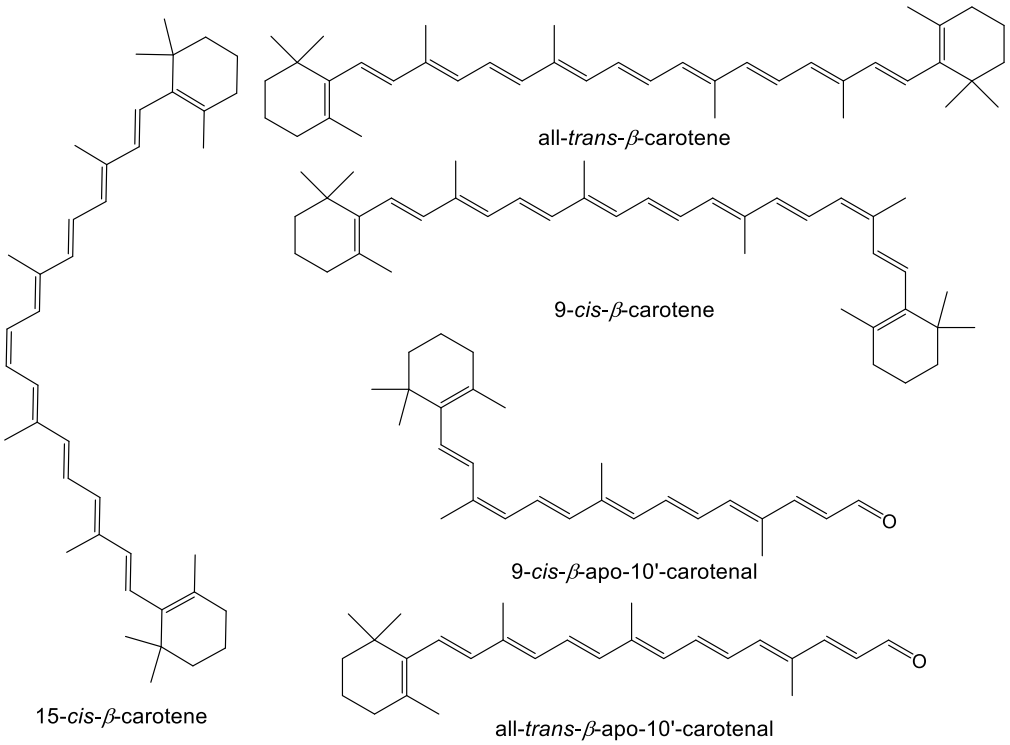


Figure 5.12 – Structures of all-*trans*-*β*-carotene, 9-*cis*-*β*-carotene, 15-*cis*-*β*-carotene, 9-*cis*-*β*-apo-10'-carotenal and all-*trans*-*β*-apo-10'-carotenal.

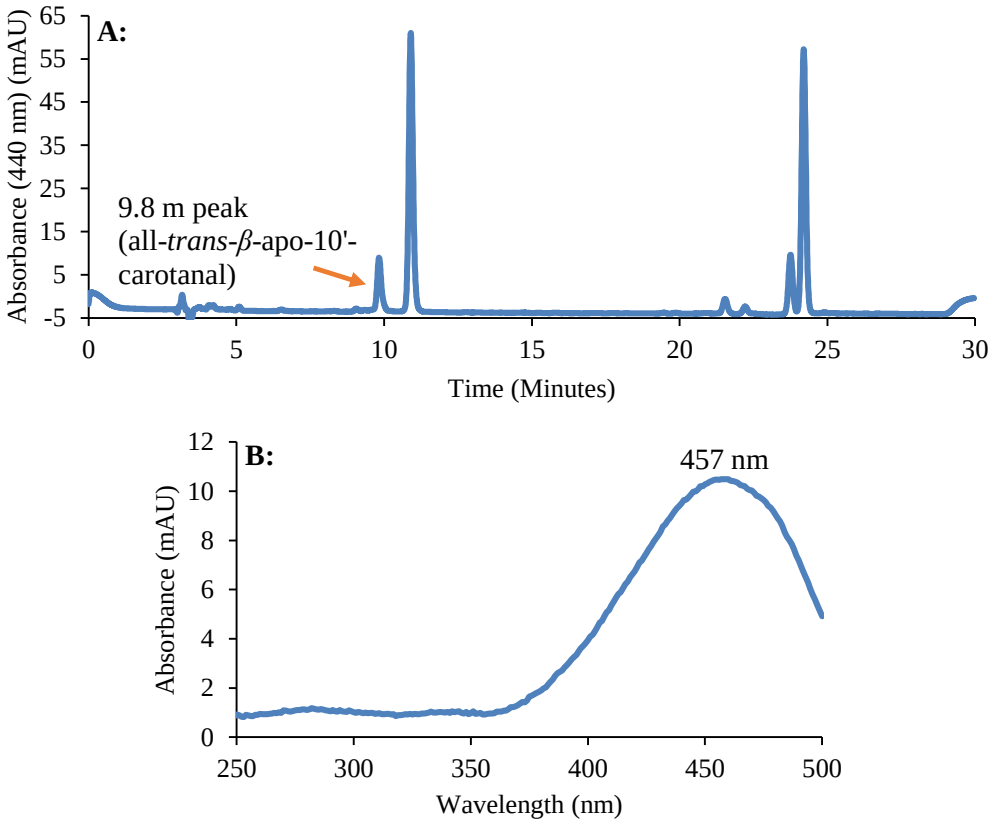


Figure 5.13 – HPLC chromatogram (**A**) showing the appearance of a peak at 9.8 minutes on following incubation of 9-*cis*-*β*-carotene with AtCCD7.2 and the UV absorption spectra (**B**) of the peak at 9.8 minutes believed to correspond to all-*trans*-*β*-apo-10'-carotenal.

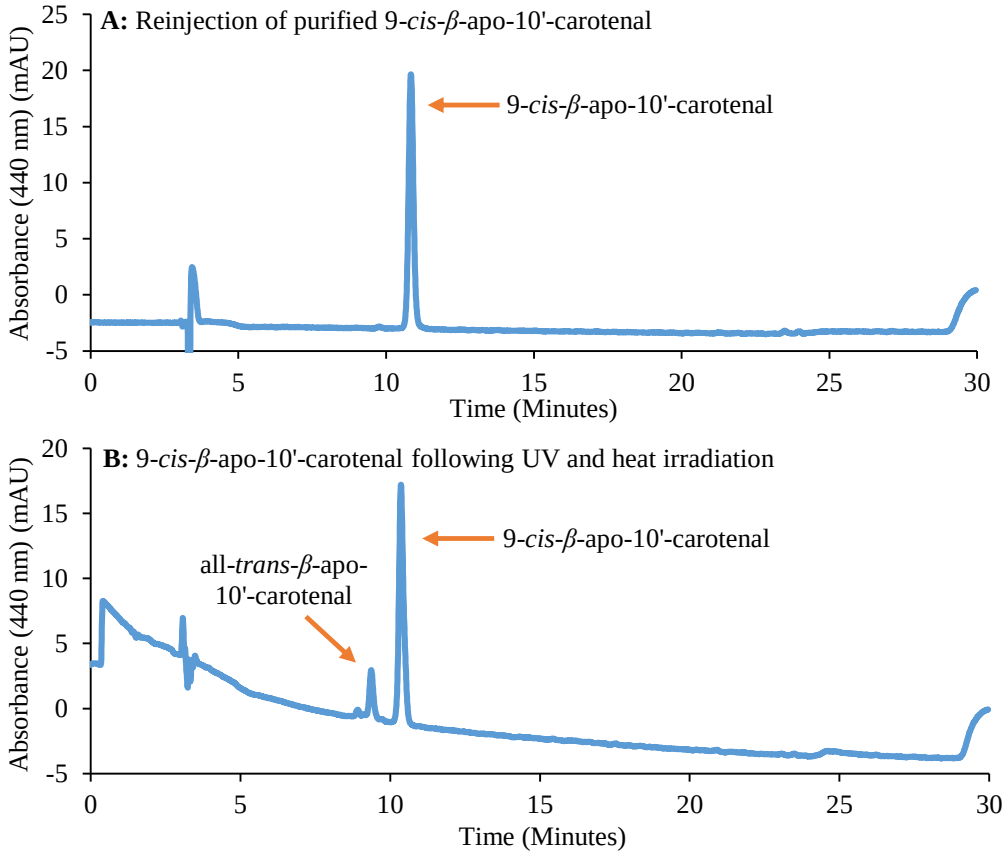


Figure 5.14 – HPLC chromatogram showing the immediate reinjection of 9-*cis*-*β*-apo-10'-carotenal following HPLC purification (**A**). Only the single species is present in the purified fraction. Following exposure to heat and light (**B**) a new peak belonging to the all-*trans* isomer is formed is formed.

peak is an isomer of 9-*cis*-*β*-apo-10'-carotenal (Figure 5.14). As such, the peak at 9.8 minutes was assigned to all-*trans*-*β*-apo-10'-carotenal. This suggests that AtCCD7.2 is predominantly selective for the 9-*cis* isomer, but is also capable of cleaving other isomers of *β*-carotene. The ability of AtCCD7.2 to cleave different *β*-carotene isomers as demonstrated here, albeit at different rates, explains the observations made by Schwartz *et al*. that CCD7 is able to cleave all-*trans*-*β*-carotene (Schwartz *et al*. 2004).

5.5 Inhibition of *A. thaliana* CCD7

Application of hydroxamic acids to *Arabidopsis thaliana in planta* results in a shoot branching phenotype, in which there is increased level of branching in the plant. Shoot branching is mediated by the hormone strigolactone, the biosynthesis of which requires CCD7

(Gomez-Roldan *et al*. 2008, Umehara *et al*. 2008). In order to investigate the biochemical basis of this *in planta* phenotype, using the HPLC assay detailed above, AtCCD7.2 was screened against the hydroxamic acid inhibitors (structures are shown in tables 2.1 to 2.5 and 3.3 to 3.7), as well as compounds D12H, D13H (synthesised by Paul Sainsbury (Sainsbury 2014)), abamineSG (provided by Syngenta) and nordihydroguaiartetic acid (NDGA, Sigma) (Figure 4.20).

All inhibitors were assayed against AtCCD7.2 at a final concentration of 100 µM except D12, which was assayed at a final concentration of 10 µM. As with NCED, the inhibitors were incubated with the AtCCD7.2 for 10 minutes prior to the addition of 9-*cis*-*β*-carotene, which was added to a final concentration of 40 µM. Assays were incubated at 25° C at pH 7.8 for 30 minutes before the assay was stopped and the products extracted as described above.

At 100 µM, none of the compounds tested showed any inhibition against AtCCD7.2 *in vitro* (Figure 5.15). In order to confirm that zero inhibition was not due to degradation of the hydroxamic acids, fresh solutions of D2 and D4 were prepared from material provided by Syngenta. These solutions at 100 µM also failed to inhibit the cleavage activity. Compound D2 was also screened against AtCCD7.2 with one hour pre-incubation, but even with this extended time no inhibition was observed. This data is surprising given that using *E. coli in*

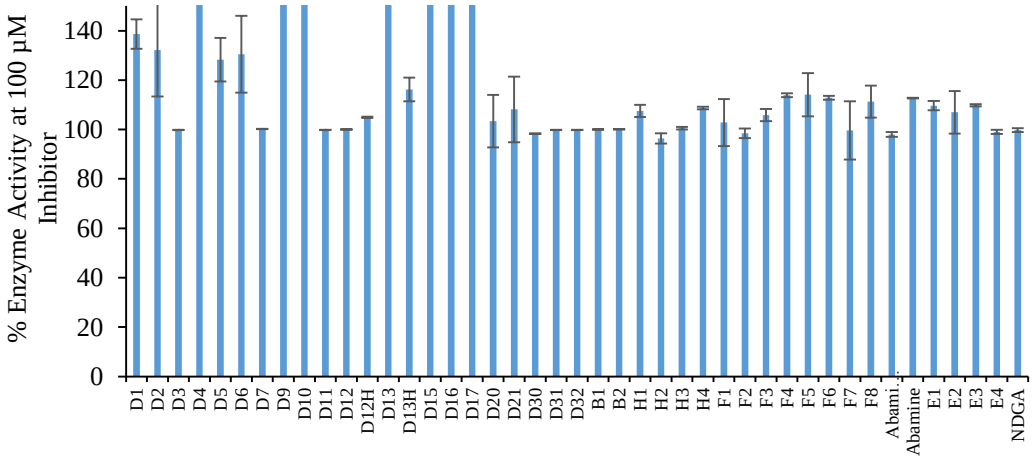


Figure 5.15 – Percentage activity of each inhibitor assayed at 100 µM against AtCCD7.2. D12 is at a concentration of 10 µM.

vivo assays, partial inhibition of AtCCD7 had been seen at 100 µM of inhibitor and CCD7 is one of the targets believed to be responsible for the shoot branching phenotype observed *in planta* with the hydroxamic acids (Sergeant *et al*. 2009). It is possible that in the *in vivo* assays the hydroxamic acids are inhibiting other enzymes and pathways in *E. coli*, disrupting the expression of either CCD7 or other genes and resulting in the apparent inhibition seen with respect to the control. The hydroxamic acids could also be enhancing the levels of carotenoids to begin with, leading to erroneous results.

In order to investigate possible causes for the lack of inhibition of AtCCD7.2, sequence alignment of CCD7 with other members of the CCD enzyme family was performed. Alignment of AtCCD7 with LeCCD1a with the ClustalΩ program (Sievers *et al*. 2011) shows that there is only 16.3% identity between the two sequences. Conversely, between LeNCED1 and LeCCD1a there is 30.7% identity between the sequences and between LeNCED1 and ZmNCED1 there is 60.6%, showing that LeNCED1, LeCCD1a and ZmNCED1 are more closely related to each other than to AtCCD7. Phylogenetic analysis of a wider selection of carotenoid cleavages show that CCD7 clusters away from other members of the family (Figure 5.16). This suggests there is an ancestral difference between CCD7 and other plant CCDs. On the basis of the phylogeny shown in Figure 5.16, CCD7 is more closely related to bacterial and cyanobacterial CCDs.

The fact that CCD7 appears to be evolutionarily distinct from other members of the CCD family and shares little identity with its other members could suggest that CCD7 is structurally quite distinct. This structural difference could result in an altered active site architecture, which renders the enzyme immune to inhibition by the hydroxamic acids.

Interestingly, CCD7 is also not inhibited by inhibitors which do not belong to the hydroxamic acid inhibitor family. Both NDGA and the abamine compounds showed no inhibition at 100 µM. Abamine and abamineSG have been shown to be reasonably weak inhibitors of CCD enzymes and as such strong inhibition against AtCCD7 was not expected.

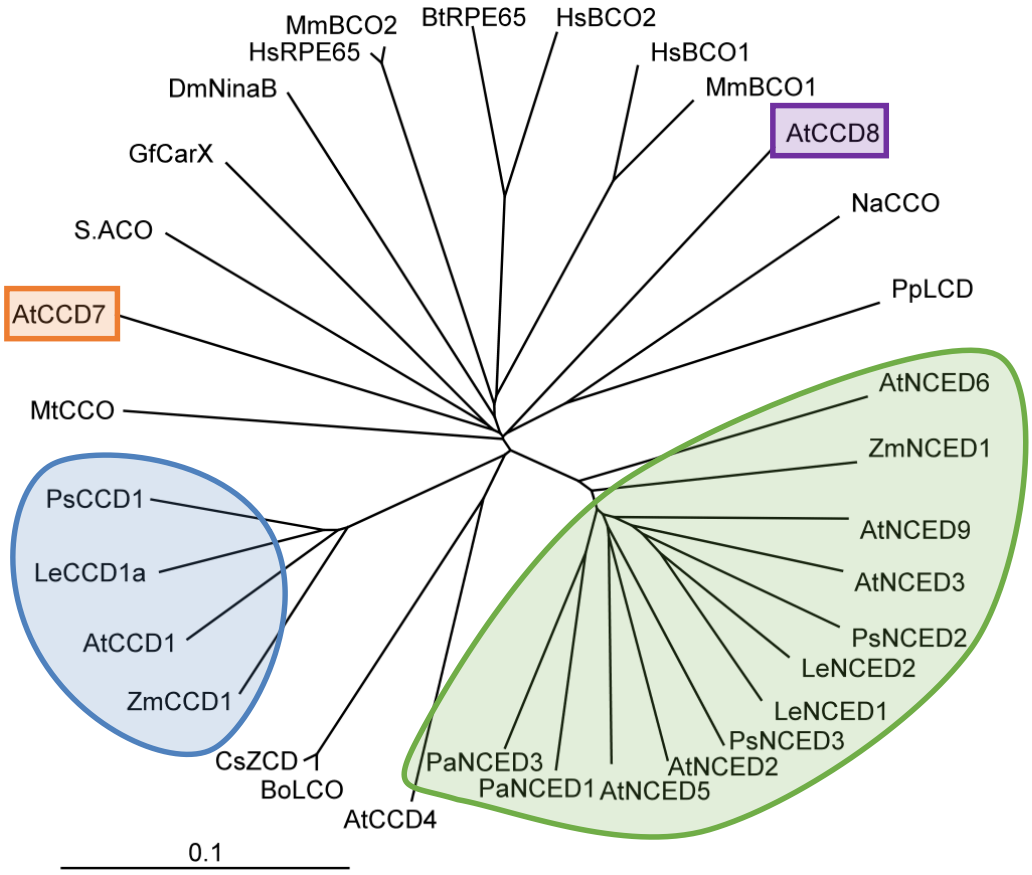


Figure 5.16 – Phylogenetic analysis of members of the carotenoid cleavage dioxygenase family following ClustalΩ alignment (See Figure 1.9). CCD7 clusters away from other members of the CCD enzyme family in plants. LCD – Lignostilbene Cleavage Dioxygenase (or LSD).

5.6 Conclusions

The CCD7 enzyme from *A. thaliana* has been shown to cleave 9-*cis*-*β*-carotene with a K_M of 8.7 µM. Additionally, AtCCD7 shows a preference for the 9-*cis* isomer of *β*-carotene over the all-*trans* isomer, confirming the results of Alder (Alder *et al*. 2012).

None of the hydroxamic acid inhibitors are effective at inhibiting the AtCCD7 cleavage reaction *in vitro*. This raises interesting questions regarding the biochemical basis for the *in planta* phenotype observed with the hydroxamic acids affecting lateral shoot branching. Neither CCD7 or D27 are inhibited by hydroxamic acids *in vitro*, and therefore are apparently not the *in planta* target. The phenotype must therefore be due to inhibition of the carotenoid cleavage dioxygenases 8 enzyme, or of another enzyme on the strigolactone biosynthesis pathway.

5.7 Future Work

Given that the *in vitro* data obtained here suggests that CCD7 is not the target of the shoot branching phenotype observed *in planta*, work remains to be done on elucidating what the *in planta* target of the hydroxamic acids is.

As with CCD1, it would be of great interest to determine the crystal structure of CCD7 and to perform detailed mechanistic experiments to help shed further light on the cleavage mechanism.

Finally, more work would be needed in order to develop an inhibitor which is selective for solely CCD7. The development of such an inhibitor would be aided by a crystal structure, preferably with bound substrate. A selective inhibitor could be useful agrochemically, but would also help to identify other processes mediated by strigolactones.

Chapter Six: Expression, Inhibition, Assays and Biochemical Characterisation of *Arabidopsis thaliana* Carotenoid Cleavage Dioxygenase 8

6.1 Introduction

Carotenoid cleavage dioxygenase 8 (CCD8) is the third enzyme on the committed pathway to strigolactone biosynthesis. CCD8 is predominantly selective for 9-*cis*-*β*-apo-10'-carotenal (produced by CCD7), performing an unusual oxidative cleavage reaction to form a compound known as carlactone, which is converted to strigolactone by the remainder of the strigolactone biosynthesis pathway (Alder *et al*. 2012). CCD8 has also been shown to cleave all-*trans*-*β*-apo-10'-carotenal, albeit at a significantly reduced rate, to form 13-apo-*β*-carotenone (Schwartz *et al*. 2004, Alder *et al*. 2008) (Figure 6.1).

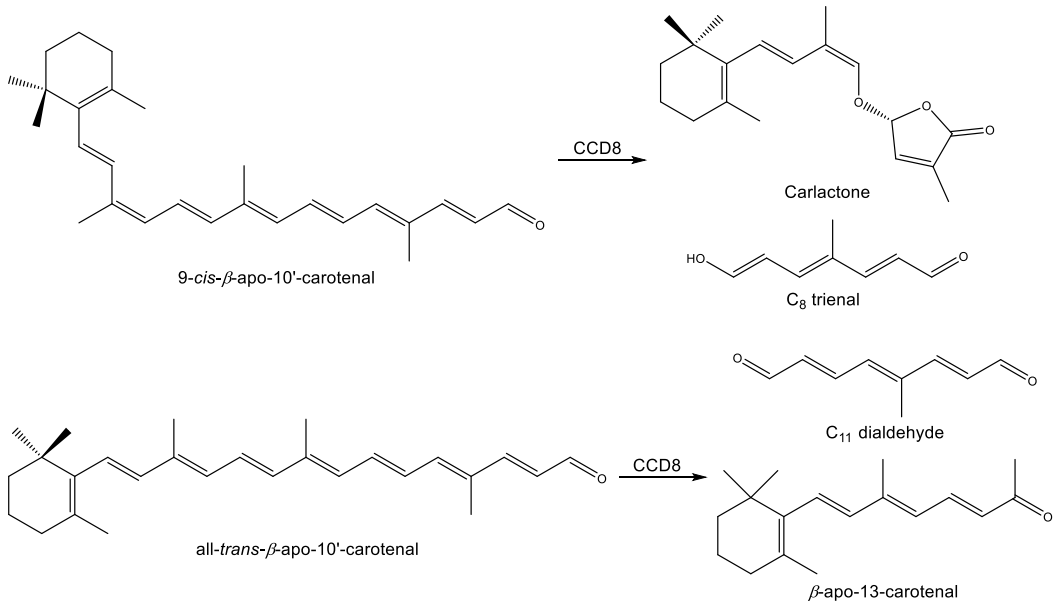


Figure 6.1 – Schematic of the reaction catalysed by CCD8.

The work detailed within this chapter describes the expression of *Arabidopsis thaliana* CCD8 and assays against hydroxamic acid carotenoid cleavage dioxygenase inhibitors in order to determine the biochemical basis of phenotypic effects on shoot branching observed in *A. thaliana in planta* on application of the hydroxamic acid inhibitors D2, D4, D5 and D6. Biochemical investigations into CCD8 are also discussed.

6.2 Expression and Assays of *A. thaliana* CCD8

Two *E. coli* constructs were obtained from Dr Andrew Thompson (University of Cranfield) containing CCD8: pGEX-4T-1-AtCCD8.1 (full length CCD8 gene with GST fusion tag) and pGEX-4T-1-AtCCD8.2 (truncated gene lacking the first 168 base pairs corresponding the chloroplast transit peptide with GST fusion tag) (CCD8 accession number: NM_119434). Both the GST-AtCCD8.1 protein (predicted molecular weight based on sequence: 90.35 kDa) and GST-AtCCD8.2 protein (predicted molecular weight based on sequence: 83.89 kDa) could be overproduced in *Escherichia coli* following induction with 1 mM IPTG at 20° C overnight and were purified by glutathione affinity chromatography (Figure 6.2). CCD8.2 was additionally cloned into the pET-151 vector and the N-His_6-AtCCCD8.2 protein (predicted molecular weight based on sequence: 61.38 kDa) could also be overproduced under the same conditions as GST-CCD8 and was purified via Ni(NTA) affinity chromatography (Figure 6.2). In the gel of GST-ATCCD8.1 (Figure 6.2A), additional bands at approximately 55 kDa and 25 kDa are due to the chaperone GroEL and cytoplasmic glutathione respectively (blue boxes).

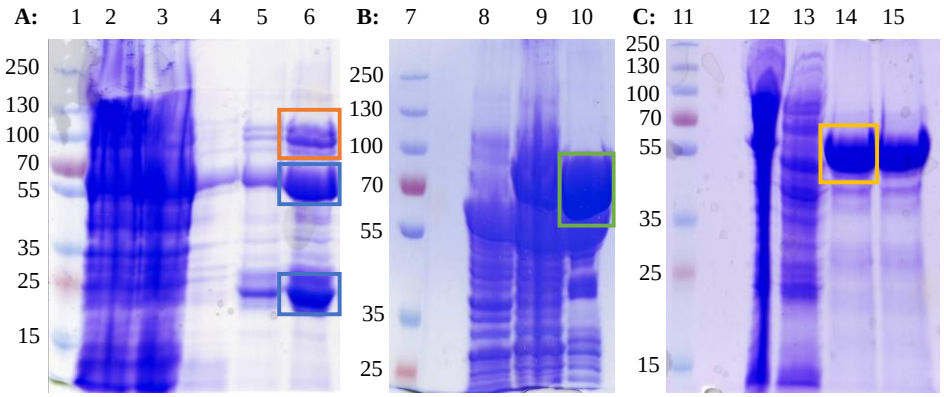


Figure 6.2 – 10% SDS-PAGE gels showing expression and purification of GST-AtCCD8.1 (**A, lanes 1-6**), GST-AtCCD8.2 (**B, lanes 7-10**) and N-His_6-AtCCD8.2 (**C, lanes 11-15**). 1 – Ladder; 2 – GST-AtCCD8.1 Total protein (TP); 3 - GST-AtCCD8.1 Total soluble protein (TSP); 4 – GST-AtCCD8.1 Flow through; 5 – GST-AtCCD8.1 Elution fraction 1; 6 – GST-AtCCD8.1 Elution fraction 2; 7 – Ladder; 8 – GST-AtCCD8.2 TP; 9 – GST-AtCCD8.2 TSP; 10 – GST-AtCCD8.2 Elution; 11 – Ladder; 12 – N-His_6-AtCCD8.2 TP; 13 – N-His_6-AtCCD8.2 TSP; 14 – N-His_6-AtCCD8.2 Elution fraction 1; 15 – N-His_6-AtCCD8.2 Elution fraction 2. Molecular weights of ladder are in kDa. Molecular weight (Mw, all calculated from sequence) of GST-AtCCD8.1: 90.35 kDa. Mw of GST-AtCCD8.2: 83.89 kDa. Mw of N-His_6-AtCCD8.2: 61.38 kDa. Mass spectrometry would be required to definitively prove that AtCCD8 has been produced.

The substrate for assays of CCD8, 9-*cis*-*β*-apo-10'-carotenal, was generated from incubations of AtCCD7.2 with 9-*cis*-*β*-carotene (generated from iodine catalysed isomerisation of all-*trans*-*β*-carotene, see Chapter 5.3). 9-*cis*-*β*-apo-10'-carotenal was purified via reversed phase HPLC using a C_{30} column and quantified using the extinction coefficient of all-*trans*-*β*-apo-8'-carotenal ($\varepsilon_{440\ nm}$ = 5.1 x 10^4 M^{-1} cm^{-1}), since no authentic standard of 9-*cis*-*β*-apo-10'-carotenal is available for the determination of an extinction coefficient. Assays of AtCCD8 contained approximately 10 µM of 9-*cis*-*β*-apo-10'-carotenal and 190 µg of protein containing overproduced CCD8.

Addition of AtCCD8 to 9-*cis*-*β*-apo-10'-carotenal was found to give a disappearance in the magnitude of the peak corresponding to 9-*cis*-*β*-apo-10'-carotenal relative to an internal standard of *β*-carotene when analysed via HPLC (Figure 6.3). Using this internal standard, the

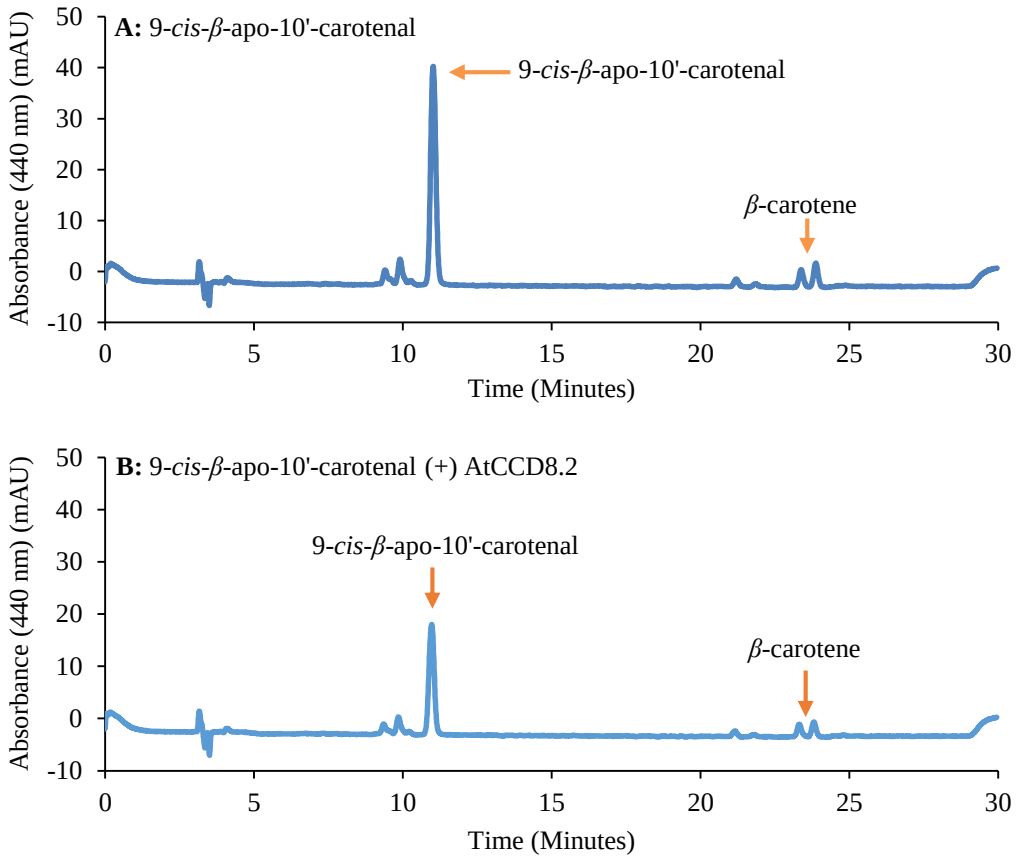


Figure 6.3 – HPLC chromatographs at 440 nm of 9-*cis*-*β*-apo-10'-carotenal in the absence (**A**, top) and presence (**B**, bottom) of N-His_6-AtCCD8.2. In the presence of N-His_6-AtCCD8.2 there is less 9-*cis*-*β*-apo-10'-carotenal with respect to the *β*-carotene in the sample (4.7:1 9-*cis*-*β*-apo-10'-carotenal : *β*-carotene versus 6.1:1).

ratio of 9-*cis*-*β*-apo-10'-carotenal to *β*-carotene was calculated, with a lower ratio indicating turnover of the 9-*cis*-*β*-apo-10'-carotenal by CCD8. Incubations of 9-*cis*-*β*-apo-10'-carotenal with both N-His_6-AtCCD8.2 and GST-AtCCD8.2 revealed a decrease in the intensity of the 9-*cis*-*β*-apo-10'-carotenal peak at 440 nm with respect to a *β*-carotene internal standard (Figure 6.3).

Activity of AtCCD8.2 could also be observed via continuous UV assays, since consumption of the 9-*cis*-*β*-apo-10'-carotenal substrate by CCD8 results in a decrease in absorbance at 430 nm. The products of the CCD8 cleavage reaction both absorb at a lower wavelength (267 nm for the carlactone and calculated 240 nm for the trienal (Williams & Fleming 1995)). As such, following incubation of 9-*cis*-*β*-apo-10'-carotenal with CCD8, a clear colour change from orange to colourless can be observed visually. Observing at 430 nm shows a decrease in absorbance on incubation of 9-*cis*-*β*-apo-10'-carotenal with CCD8, which can be followed on the plate reader (Figure 6.4). Control experiments demonstrated the change in absorbance was due to the activity of CCD8 (Figure 6.5). No activity was observed in the presence of boiled AtCCD8.2 (Figure 6.5A). Activity was observed in the presence of EDTA (Figure 6.5B), which suggests the active site iron is strongly coordinated by the enzyme. Additionally, the rate of the cleavage reaction was dependent upon the concentration of AtCCD8.2 present (Figure 6.5C).

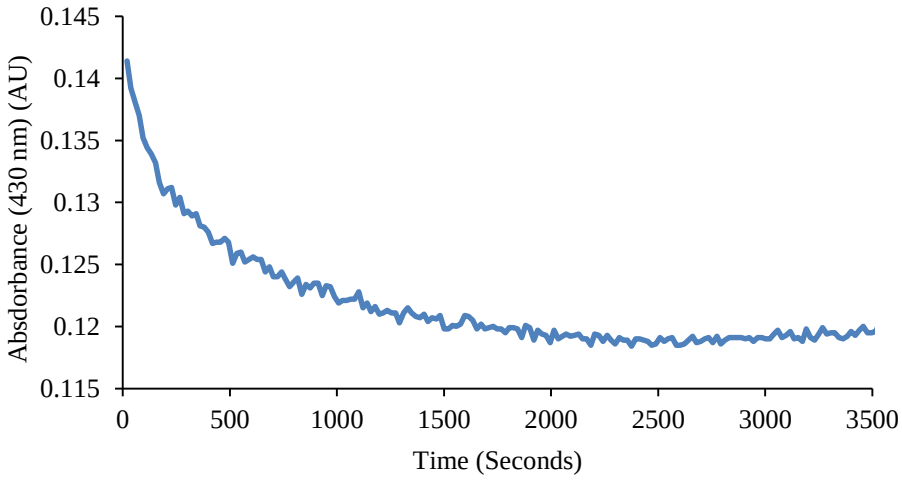


Figure 6.4 – **A:** Absorbance change at 430 nm on incubation of 9-*cis*-*β*-apo-10'-carotenal with AtCCD8.2.

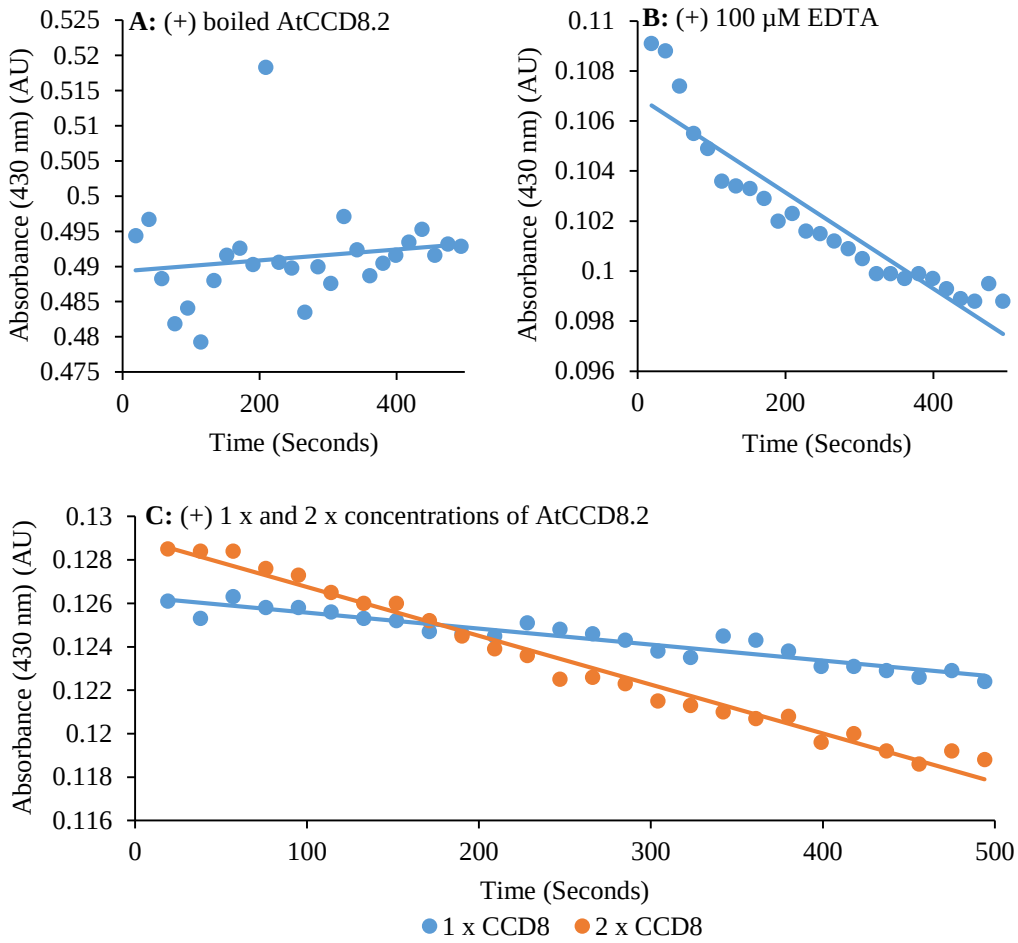


Figure 6.5 – UV chromatographs of AtCCD8 control reactions. **A:** 9-*cis*-*β*-apo-10'-carotenal with boiled AtCCD8.2. **B:** 9-*cis*-*β*-apo-10'-carotenal with AtCCD8.2 and 100 µM EDTA. **C:** 9-*cis*-*β*-apo-10'-carotenal with 1 x and 2 x concentrations of N-His_6-AtCCD8.2.

Following incubations of GST-AtCCD8.1, GST-AtCCD8.2 and N-His_6-AtCCD8.2 with 9-*cis*-*β*-apo-10'-carotenal at 25° C and analysis via HPLC it was not possible to detect a new peak at 267 nm which would correspond to the carlactone product using the same analysis method as was used for CCD7 and D27 (Alder *et al*. 2012) (Figure 6.6). The product of the CCD8 reaction, carlactone, is an enol ether (Figure 6.1), which is likely to be unstable, and can readily undergo hydrolysis in mildly acidic aqueous conditions (Figure 6.7). It is likely that during the extraction process and analysis, the carlactone is being hydrolysed by water in the presence of a protic solvent such as methanol. Breakdown of the carlactone is likely to produce an aldehyde and a hemiacetal lactone. The C_8 trienal product was also not detected. These products are unlikely to be retained on a hydrophobic C_{30} column, most likely eluting immediately after the dead volume. These small compounds are also likely volatile (a sweet

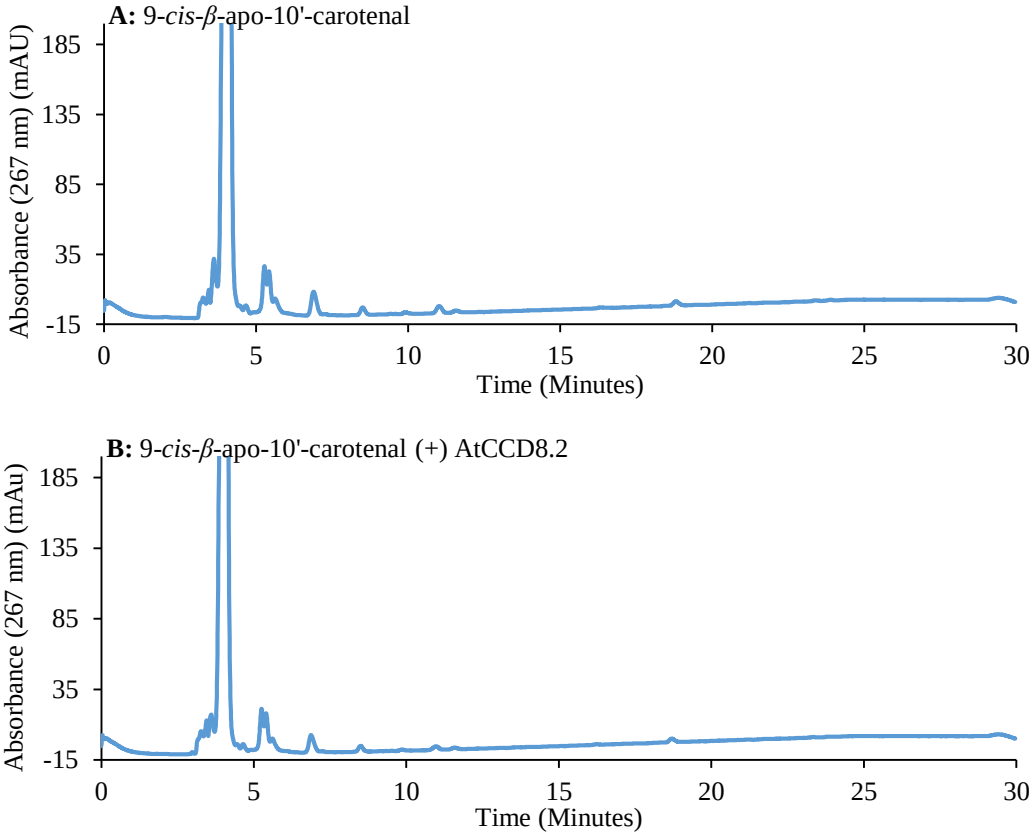


Figure 6.6 – Comparison of HPLC chromatographs at 267 nm of 9-*cis*-*β*-apo-10'-carotenal in the absence (**A**, top) and presence (**B**, bottom) of AtCCD8.2 using the CCD7/D27 analysis method. No new peaks are present.

smell can be detected following incubations of AtCCD8 with 9-*cis*-*β*-apo-10'-carotenal), and so consequentially may be lost during sample preparation. The retention of carlactone on the C_{30} column may also be weak and extractions from the assay reaction with more polar solvents such as ethyl acetate also failed to produce a peak corresponding to the carotenal.

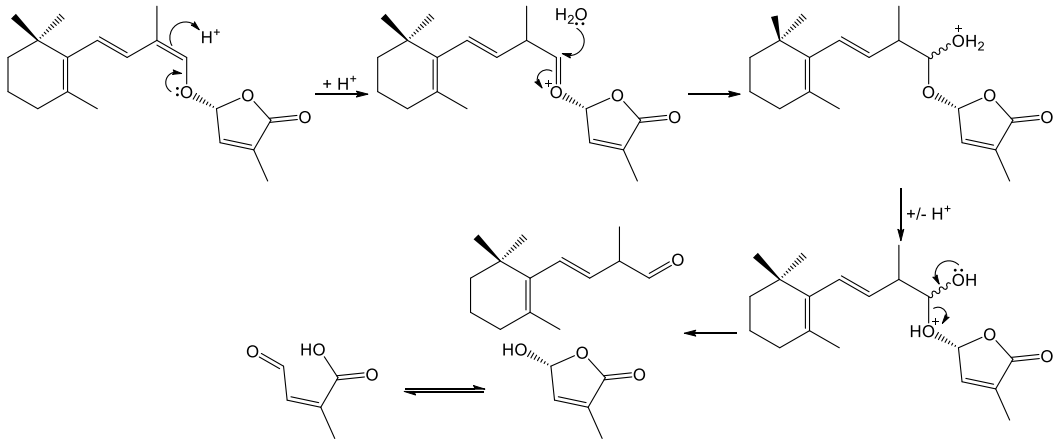


Figure 6.7 – Mechanism of the breakdown of the enol ether carlactone in mild acidic / aqueous conditions. The lactone product is possibly stabilised by the *cis* double bond.

Further analysis was performed using a less hydrophobic C_{18} column, with a mobile phase of water, methanol and tertbutylmethylether (Alder *et al*. 2012). Via this method a small peak at 15.2 minutes at 267 nm was detected in incubations of AtCCD8 with 9-*cis*-*β*-apo-10'-carotenal, as reported by Alder *et al*. (Figure 6.8). The identity of the peaks at

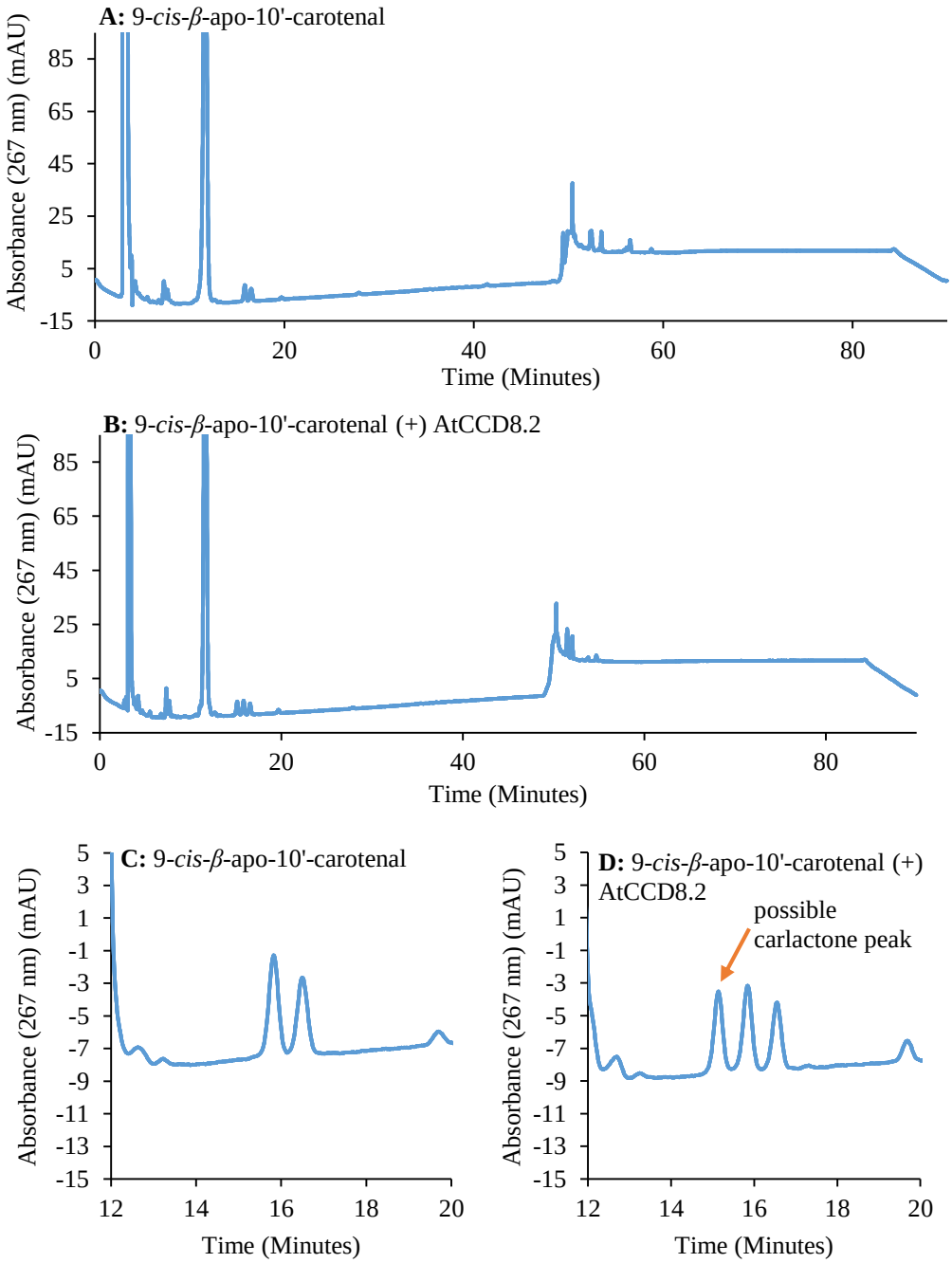


Figure 6.8 – HPLC chromatograph of incubation of 9-*cis*-*β*-apo-10'-carotenal in the absence (**A** and **C**) and presence (**B** and **D**) of N-His_6-AtCCD8.2. **B** and **D** show enhanced view of the region between 12 and 20 minutes, where a new peak is present following incubation of 9-*cis*-*β*-apo-10'-carotenal with AtCCD8.2.

15.5 minutes and 16.5 minutes is unknown. Analysis of the UV spectra of this peak at 15.2m showed a chromatogram identical to that reported by Alder *et al.* for carlactone (Figure 6.9). However, it was not possible to perform liquid chromatography mass spectrometry (LC-MS) directly on this sample due to the nature of the solvents employed. Analysis via low resolution mass spectrometry did not reveal a peak which would correspond to carlactone. This molecule appears to be very challenging to detect via mass spectrometry.

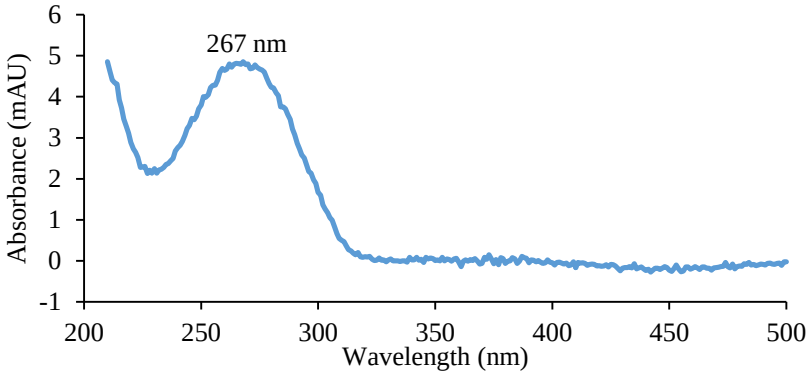


Figure 6.9 – UV chromatogram of the peak at 15.2 minutes from Figure 6.6, possibly belonging to carlactone.

In order to perform LC-MS, a mobile phase of solely methanol / water was employed. However, using this method it was not possible to discern a new peak in the UV chromatogram which would correspond to the carlactone product, and no signals at *m/z* = 303 could be detected in the extracted ion chromatogram, which would correspond to the $[M+H]^+$ signal of carlactone. This lack of signal strongly indicates that the carlactone is breaking down during the extraction procedure as suggested above. However, sodium borohydride treatment of the organic extract of the CCD8 reaction (Figure 6.10), followed by analysis via low resolution ESI mass spectrometry revealed peaks with masses of *m/z* = 209 (the reduced aldehyde breakdown product $[M+H]^+$ peak), *m/z* = 143 (the reduced trienal product $[M+H]^+$ peak) and *m/z* = 379 (unreacted 9-*cis*-*β*-apo-10'-carotenal $[M+H]^+$ peak). Although it was not possible to obtain high resolution spectra, these data indicate that CCD8 is indeed converting 9-*cis*-*β*-apo-10'-carotenal to carlactone (low resolution spectra are shown in appendix 10.6)

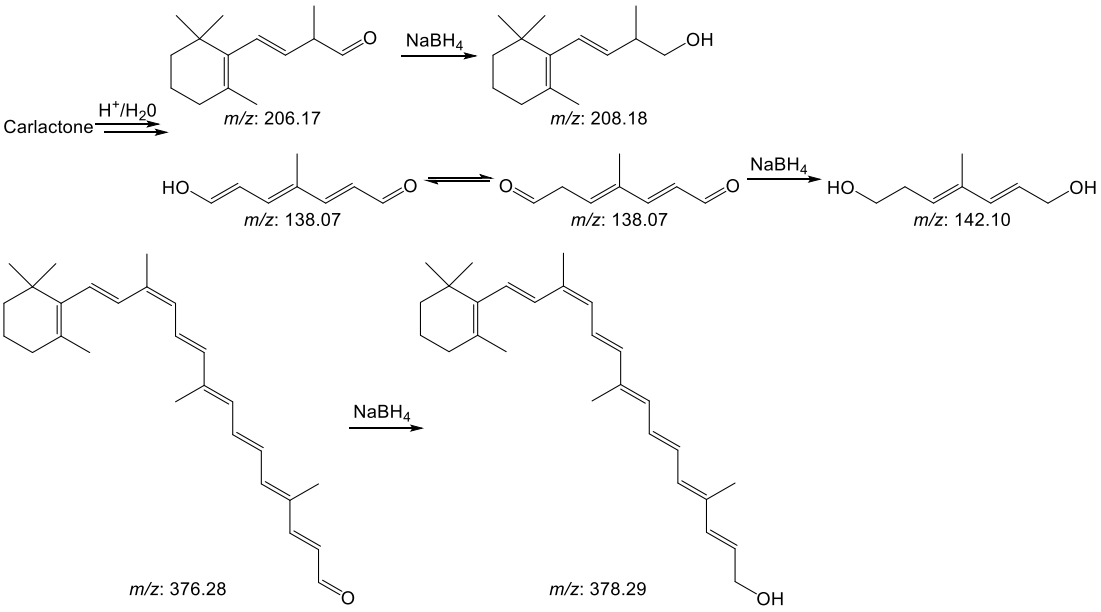


Figure 6.10 – *m/z* values of CCD8 reaction products, breakdown products and reduced derivatives.

Consequentially it can be concluded that AtCCD8 is catalysing the cleavage of 9-*cis*-*β*-apo-10-carotenal, as reported by Alder *et al.*, however, it is not possible to definitively confirm the presence of the carlactone product.

6.3 Crystallisation of *A. thaliana* CCD8.2

N-His_6-AtCCD8.2 was purified via nickel affinity chromatography for the purposes of crystallisation of the AtCCD8 protein. Purified N-His_6-AtCCD8.2 was added to a panel of 100 crystallisation buffers provided by Dr Dean Rea (University of Warwick, Life Sciences) and allowed to crystallise over a period of 7 days. Crystals were observed for N-His_6-AtCCD8.2 in the presence and absence of addition iron (II) sulfate (1 µM) (Figure 6.11) for the buffers Morpheus C2 (small cubes) (10 % w/v PEG 8000, 20 % v/v ethylene glycol,

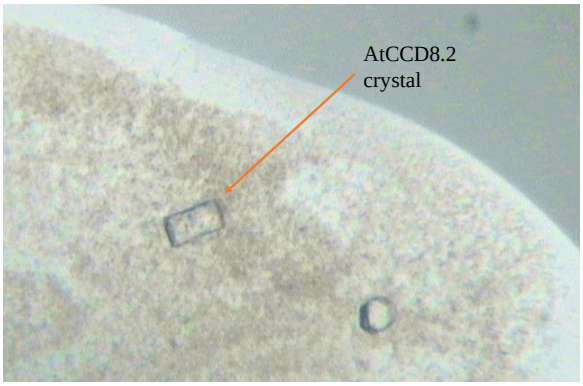


Figure 6.11 – Image of AtCCD8.2 protein crystal produced from the Morpheus C6 buffer.

0.03 M of each, 0.1 M MES/Imidazole pH 6.5), Morpheus C6 (single rocks) (10 % w/v PEG 8000, 20 % v/v ethylene glycol, 0.03 M of each NPS, 0.1 M MOPS/HEPES-Na pH 7.5) and Morpheus C10 (rod clusters) (10 % w/v PEG 8000, 20 % v/v ethylene glycol, 0.03 M of each NPS, 0.1 M Bicine/Trizma base pH 8.5). Due to time constraints, it was not possible to perform further experiments towards obtaining a crystal structure of AtCCD8.

6.4 Biochemical Characterisation of *A. thaliana* CCD8

CCD8 is proposed to perform an unusual cleavage reaction, inserting three equivalents of oxygen, which are all believed to be derived from molecular oxygen, in a single reaction, forming carlactone. Currently, there is no kinetic, biochemical or mechanistic data on CCD8 in the literature, nor is there a crystal structure of CCD8 available to help probe the mechanism.

Attempts were made to determine a K_M for CCD8 for the substrate 9-*cis*-*β*-apo-10'-carotenal. However, it was observed that at a concentration of substrate above 15 µM there was a decrease in the rate of the CCD8 reaction (Table 4.1). The rate of the reaction at various substrate concentrations was recorded and modelled using analysis software (GraphPad) and was found to be consistent with a model of substrate inhibition (Figure 6.12). This possibly arises through the enzyme binding a second equivalent of substrate in a conformation which is detrimental to the activity of the enzyme. A K_M of 9.2 µM was calculated by the software and specific activity of 2.48 x 10^{-3} units (µmol s^{-1}) mg^{-1} was determined for N-His_6-AtCCD8.2. A k_{cat} of 0.18 s^{-1} was also determined for N-His_6-

Substrate Concentration (µM)	Reaction Rate (nmol min^{-1})
1	0.0
3	0.0022
5	0.0163
7	0.0188
15	0.0147
30	0.0
45	0.0
60	0.0

Table 6.1 – Comparison of the rate of the CCD8 reaction at various substrate concentrations.

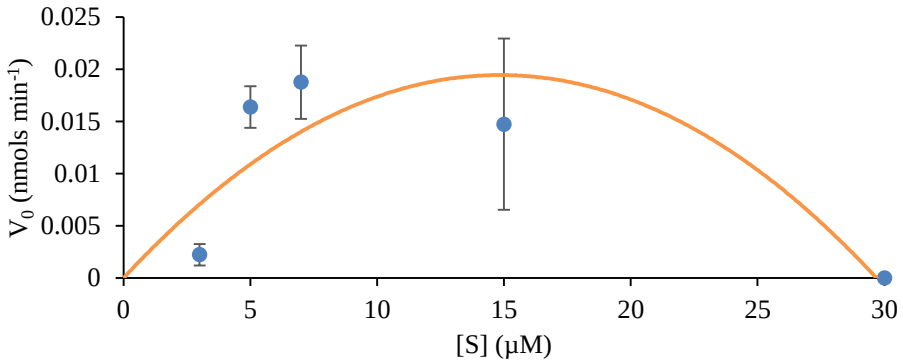


Figure 6.12 – Modelling of AtCCD8.2 rate data using a model of substrate inhibition by GraphPad.

AtCCD8.2. It should be noted that the values of k_{cat}, K_M and the specific activity are based on the assumption that 80% of protein present was that of N-His_6-AtCCD8.2, and that concentration of substrate was determined using the extinction coefficient of all-*trans*-*β*-apo-8'-carotenal. A k_{cat}/K_M value of 0.0199 s^{-1} μM^{-1} was also determined for N-His_6-AtCCD8.2. This kinetic data suggests that CCD8 is a slow enzyme that does not bind the 9-*cis*-*β*-apo-10'-carotenal substrate tightly. This is consistent with the fact that CCD8 is catalysing a complex cleavage reaction.

In order to investigate the mechanism of AtCCD8, stopped-flow experiments were performed on a visit to the group of Prof. Nigel Scrutton (University of Manchester) using a

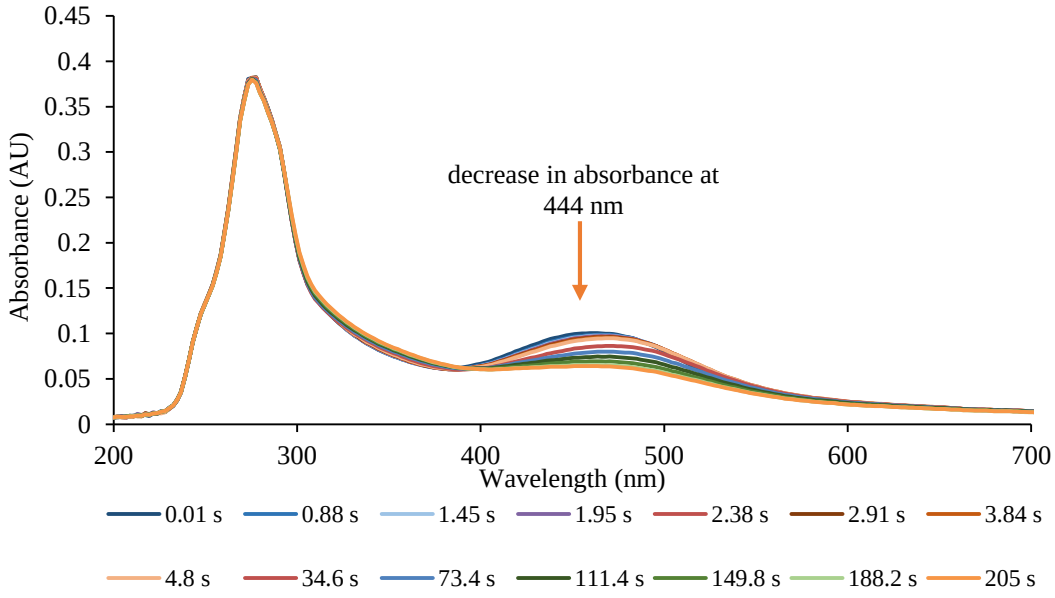


Figure 6.13 – Absorbance change observed over 200 s from stopped flow reaction of 11 µM AtCCD8.2 with 11 µM 9-*cis*-*β*-apo-10'-carotenal in 100 mM HEPES buffer pH 7.8 with 1 mM TCEP. A decrease in absorbance can be observed at 444 nm.

1:1 molar ratio of AtCCD8.2 to 9-*cis*-*β*-apo-10'-carotenal (Figure 6.13). Following the reaction over a period of 200 s using a diode array detector in 100 mM HEPES buffer pH 7.8 with 1 mM TCEP, a decrease in absorbance at 440 nm can be observed, which represents consumption of the 9-*cis*-*β*-apo-10'-carotenal substrate by CCD8 (Figure 6.13). Further kinetic experiments were performed over a time period of 40 s at a fixed wavelength of 440 nm (Figure 6.14). Analysis of the data revealed that decay of the 9-*cis*-*β*-apo-10'-carotenal at 440 nm could be fitted well to a second order exponential (Figure 6.14). This suggests that the mechanism of cleavage by CCD8 is a two-step process (Figure 6.15A).

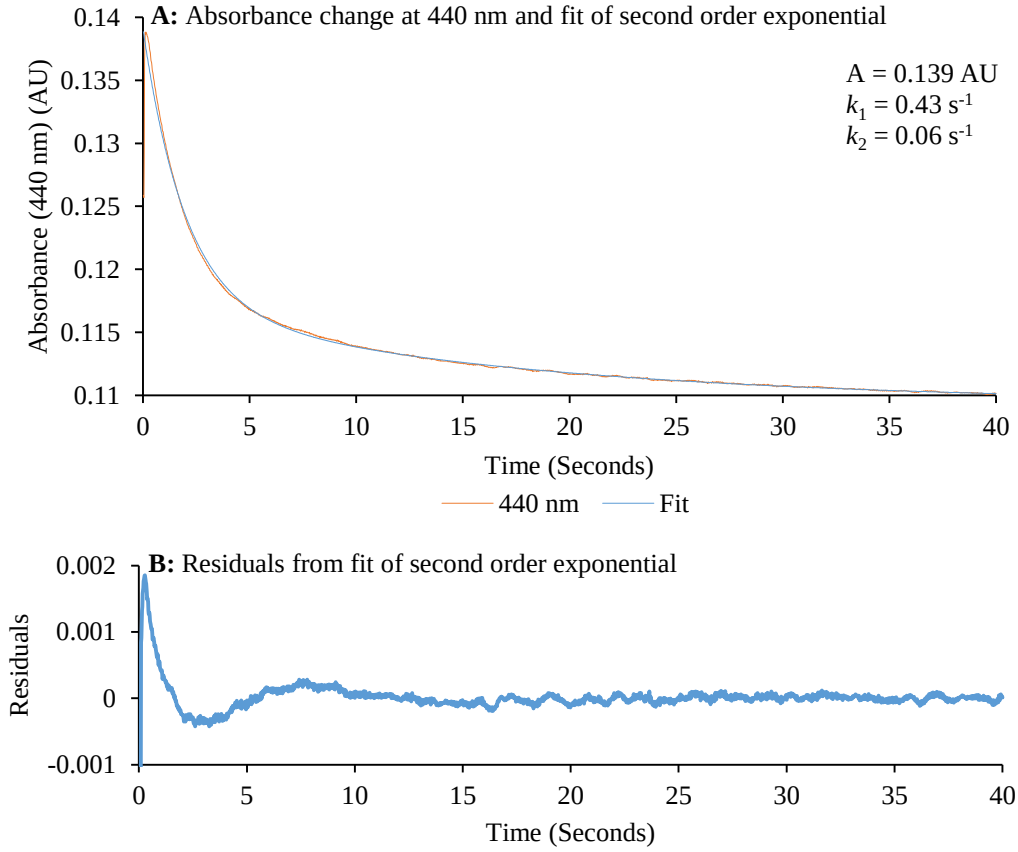


Figure 6.14 – **A:** Absorbance change observed over 40 s at 440 nm for reaction of 11 µM AtCCD8.2 with 11 µM 9-*cis*-*β*-apo-10'-carotenal in 100 mM HEPES buffer pH 7.8 with 1 mM TCEP (orange) and fit to a second order exponential (blue). **B:** Residuals from the fitting of second order exponential to absorbance change at 440 nm.

There is also a slight increase in absorbance at approximately 350 nm during the reaction, which could represent the formation of a conjugated intermediate during the reaction process (Figure 6.13). However, it was not possible to fit the change at 350 nm to either a single or double exponential. No absorbance change could be detected at 267 nm, which would

correspond to the carlactone product, but this may be hidden in the presence of a high protein concentration. Hence the two step mechanism implies the presence of a less UV absorbent intermediate in the range of 5-15 seconds (Figure 6.15B).

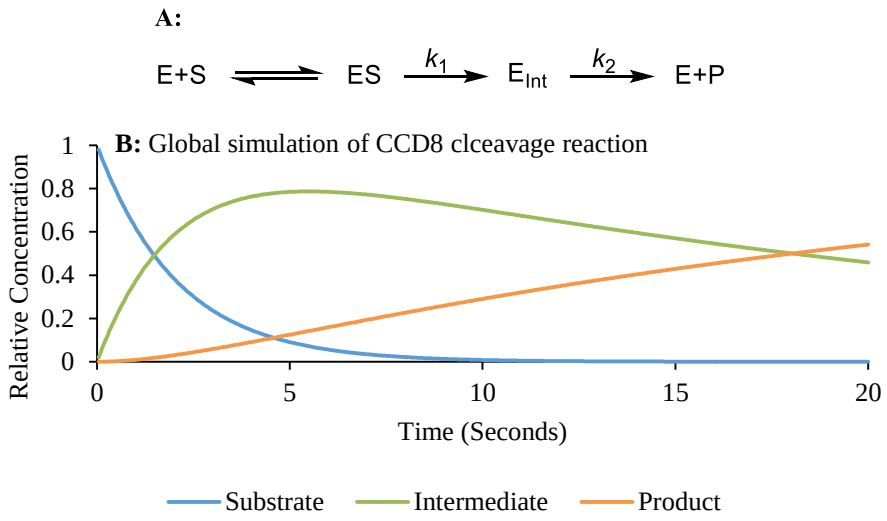


Figure 6.15 – **A:** Possible reaction scheme for the CCD8 cleavage reaction involving an enzyme bound intermediate. **B:** Global simulation of CCD8 cleavage reaction showing relative changes in the concentrations of substrate, intermediate and product versus time.

A pH rate profile was performed to determine if any acid-base catalysis was occurring during the CCD8 cleavage reaction (Figure 6.16). The pH rate profile analysis reveals that there is a sharp decrease in the rate of the CCD8 catalysed reaction below pH 6.0 and above pH 8.0. The fall in rate at low pH could be due to protonation of a basic residue, such as histidine (pK_a = 6-8), glutamic acid or aspartic acid (pK_a = 4-6) (Bugg 2012). Conversely the fall in rate at high pH could be due to the deprotonation of an acidic residue,

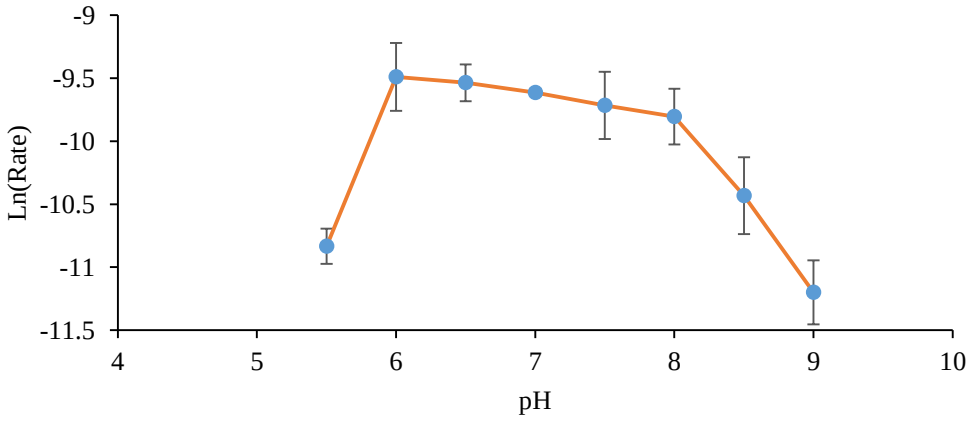


Figure 6.16 – pH rate profile of the AtCCD8.2 catalysed cleavage of 9-*cis*-*β*-apo-10'-carotenal.

such as cysteine (pK_a = 7.5-9), lysine (pK_a = 8.5-10), histidine or even iron bound water (Fe-OH_2). Consequentially, the pH rate profile data strongly suggests that acid-base catalysis is required for cleavage of 9-*cis*-*β*-apo-10'-carotenal by CCD8.

Based on the data of the pH rate profile, group specific reagents were chosen in order to determine the nature of residues which may be involved in catalysis. It should be noted that group specific reagents will also modify non-catalytic residues, which could perturb the activity of the enzyme. CCD8 was incubated for 20 minutes in the presence of a series of group specific reagents (Table 6.2) before being assayed as usual. It was observed that CCD8 was inhibited by iodoacetamide, *N*-ethyl maleimide (NEM), diethyl pyrocarbonate and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC). The loss of enzyme activity in the presence of these chemicals suggests the presence of cysteine (iodoacetamide and NEM), histidine (diethyl pyrocarbonate) and aspartic acid or glutamic acid (EDC) residues in or near the enzyme active site. It is known that within the CCD enzyme family there are semi-conserved second shell glutamic acid residues which could well be participating in catalysis (Harrison & Bugg 2013, see also Figure 1.9). Histidine could be acting either as an acid or a base, or the inhibitory effect could be a result of modification of the histidine residues required to coordinate the active site iron, resulting in a loss of ability to coordinate the iron. Cysteine could be either acting as an active site acid or in a redox capacity. No loss of activity was observed in the presence of succinic anhydride, a group specific reagent for lysine residues.

Reagent	Target Residue	% Activity after 20 minutes
Iodoacetamide	Cysteine	< 1%
N-ethyl maleimide	Cysteine	< 1%
Diethyl pyrocarbonate	Histidine	< 1%
1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide	Aspartic acid / Glutamic acid	< 1%
Succinic anhydride	Lysine	> 95%

Table 6.2 – Comparison of effects observed for different group specific reagents on the AtCCD8 enzyme reaction.

6.5 Discussion of Possible CCD8 Cleavage Mechanisms

Understanding the CCD8 cleavage mechanism is complicated by the fact that CCD8 performs an unusual cleavage reaction, apparently cleaving two C-C bonds rather than two C=C bonds (although the polyene backbone is a conjugated system, hence electron density is spread equally amongst the carbon atoms). As a result, CCD8 must proceed through a non-standard mechanism, i.e. one unlike the other members of the CCD enzyme family (such as those discussed in Chapter 1.6). Experiments with group specific reagents suggest the presence of an active site cysteine residue, which could act as a nucleophile, and has the

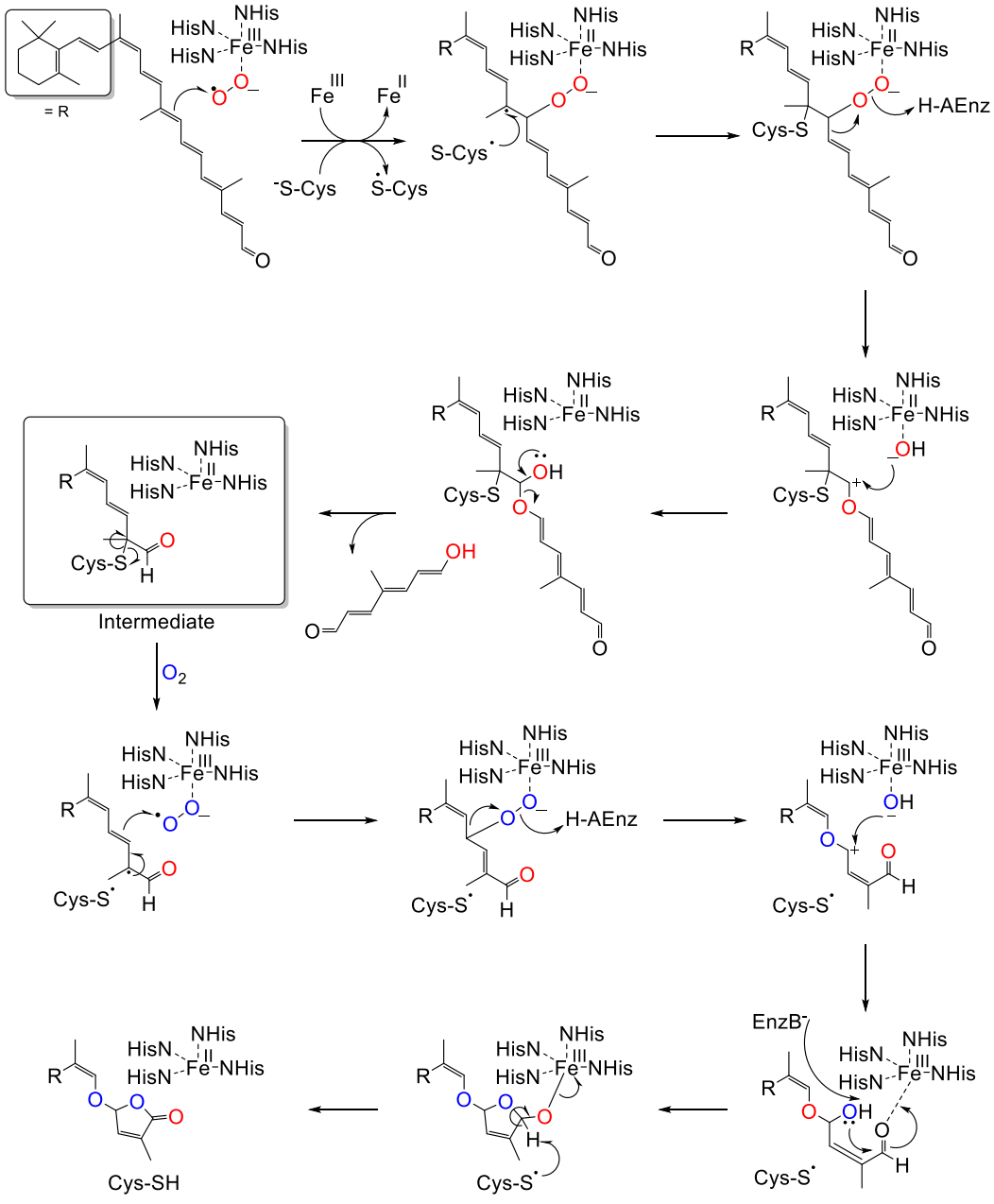


Figure 6.17 – Possible mechanism for the CCD8 cleavage reaction.

potential to form a stable radical intermediate (Stubbe & van der Donk 1998). A possible mechanism for CCD8 is shown in Figure 6.17.

In this mechanism, following formation of the iron-peroxo intermediate by iron (II) a stabilized tertiary radical is formed. One electron transfer from cysteine to iron (III) would then occur, creating a cysteinyl radical, which attacks the substrate radical at C13. The less hindered group of the substrate is then able to migrate via a Criegee rearrangement (Bugg 2003), creating an enol ether and secondary cation, which is quenched by the iron bound hydroxyl. The hydroxyl group can then eliminate the trienal, which is a good leaving group, leaving an enzyme bound covalent intermediate. Homolytic cleavage of the S-C bond forms a highly stabilized tertiary radical, which allows a second iron-peroxo intermediate to form at C11. This time, the more electron rich group is able to migrate in the Criegee rearrangement and again the secondary cation is quenched by the iron bound hydroxyl. Base catalysed cyclisation onto the aldehyde then creates a five membered lactol. Finally, oxidation of the lactol by the active site cysteine radical and iron (III) restores the active site iron to the iron (II) resting state and releases the carlactone product.

Although there is no definitive proof for the existence of this mechanism, it is consistent with the available evidence; that is a two-step acid-base catalysed mechanism involving at least three different active site residues. The sequence of cleavages also makes sense, given that a cyclisation reaction is required to form the lactone ring, which would first require aldehyde formation. Analysis of stopped flow data reveals that the rate constants for the first and second reaction steps are 0.43 s^{-1} and 0.06 s^{-1} respectively, in close agreement to the value of 0.18 s^{-1} calculated from a continuous UV assay (Section 6.4). As such, this equates to a turnover time of 2.3 s and 16 s respectively for the first and second steps. A slow second reaction is consistent with this mechanism, as the second half of the reaction requires homolytic cleavage of the reasonably stable C-S bond. Sequence alignment of CCD8 homologs shows that there are a number of conserved cysteine residues in the CCD8 enzyme family (Cys_{217}, Cys_{294}, Cys_{394}, Cys_{395}, Cys_{467} and Cys_{479} in *A. thaliana* CCD8) (Figure 6.18).

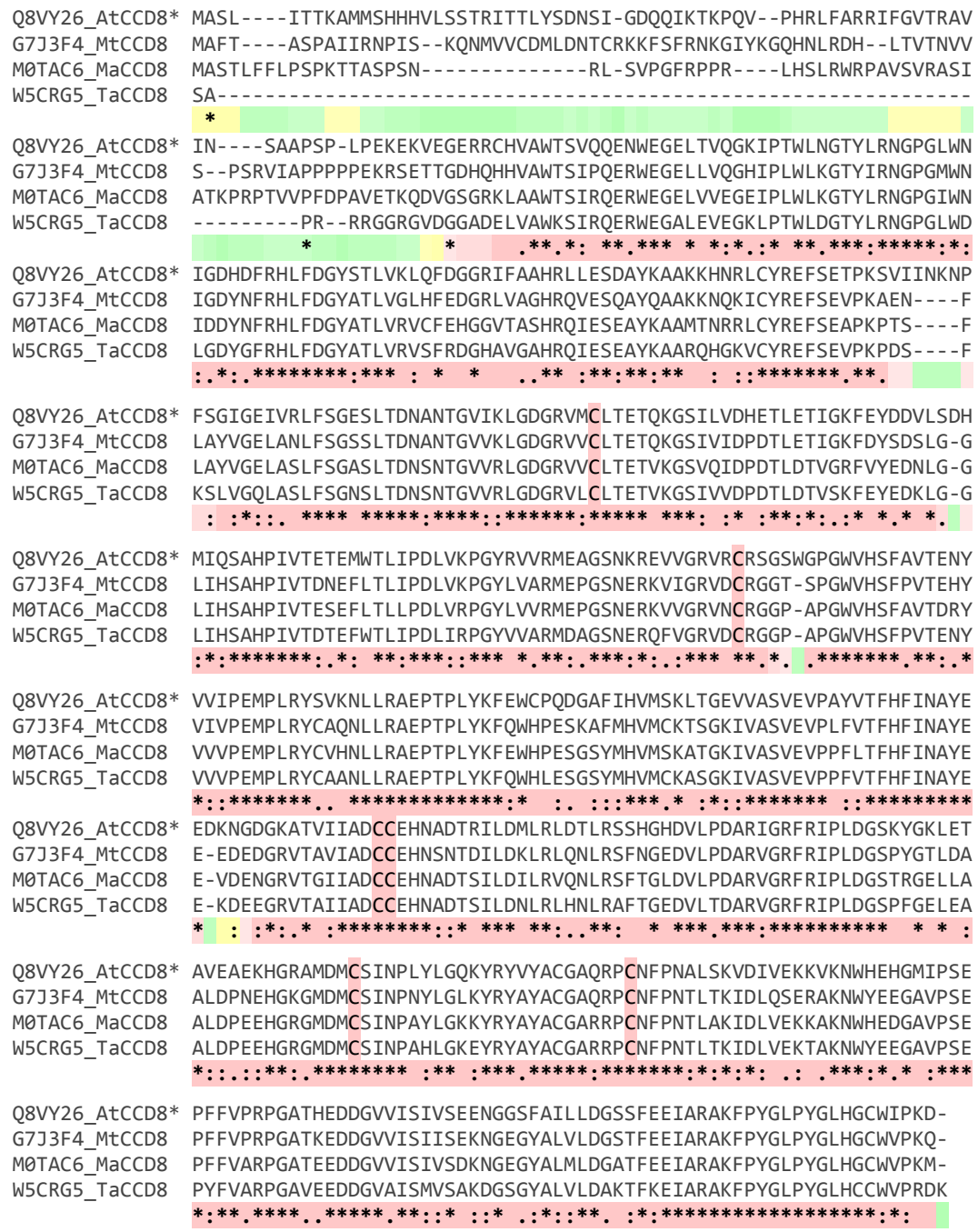


Figure 6.18 – Multiple T-Coffee sequence alignment of CCD8 against selected random hits from a BLAST search. Conserved cysteine residues are highlighted in red (based on alignment to top 50 hits in BLAST search). At – *Arabidopsis thaliana;* Mt – *Medicago truncatula* (barrel clover); Ma – *Musa acuminata* (wild banana); Ta – *Triticum aestivum* (wheat). * refers to functionally characterised enzyme.

Another CCD8 mechanism has been proposed by Alder *et al*. (Figure 6.19) (Alder *et al*. 2012). In this mechanism, the 9-*cis*-*β*-apo-10'-carotenal substrate undergoes a double isomerisation reaction to form 9,11,13-tri-*cis*-*β*-apo-10'-carotenel, which then binds to the

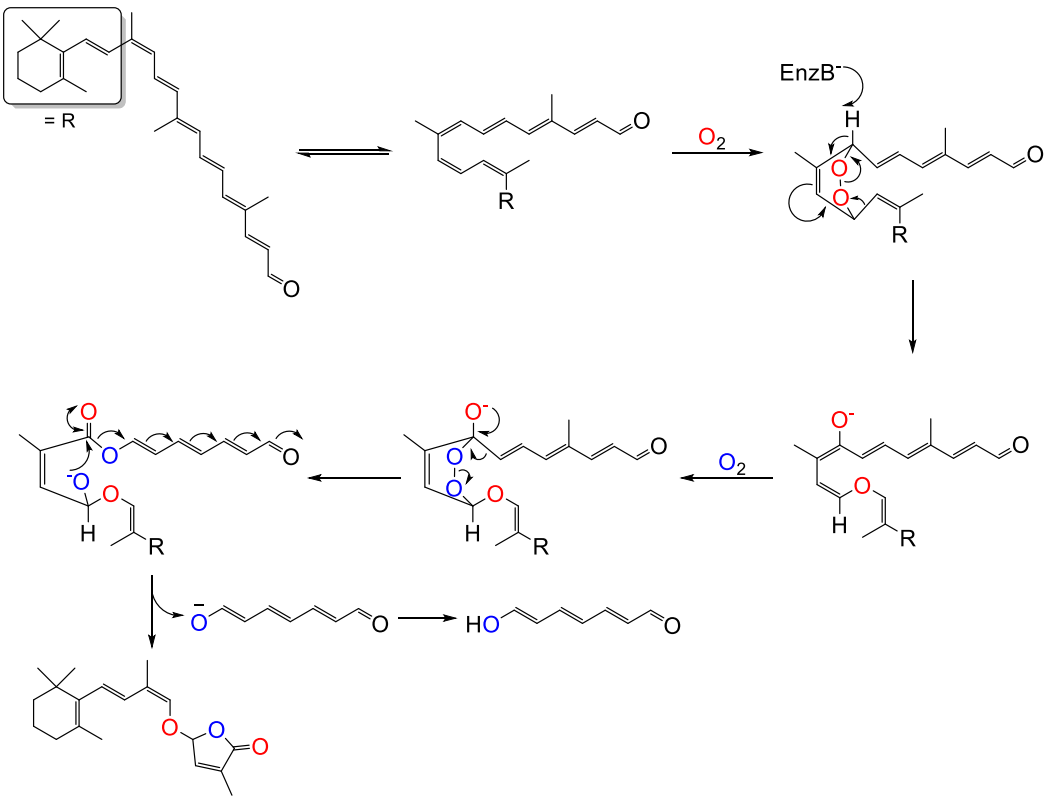


Figure 6.19 – Proposed mechanism by Alder *et al*. for the formation of carlactone by AtCCD8. Figure from Alder *et al*. 2012.

active site in a manner similar to cyclooxygenase. Molecular oxygen is then added across two double bonds to form a dioxane intermediate, which undergoes base catalysed rearrangement to form an enol ether and an oxygen anion intermediate. A second equivalent of oxygen is then added across the same pair of double bonds as before. This structure then rearranges to form another oxygen anion, which can then cyclise onto the resulting ester, eliminating the trienal and forming carlactone.

There are several problems with this mechanism, however. Although it is a direct route, dioxane formation seems unlikely, as CCD8 bears no sequence similarity to cyclooxygenases (van der Donk *et al*. 2002). Moreover, it is not clear why one isomerase, D27, would evolve on the pathway to perform one isomerisation, only for CCD8 to carry out two further isomerisation reactions. It is likely that if CCD8 were to catalyse an isomerisation as proposed in Figure 6.19, it would require an iron-oxo intermediate, as discussed in Chapter 4. Such an intermediate would likely form cleavage products, which are not observed.

Additionally, it is likely that CCD8 would also be able to isomerise the all-*trans*-*β*-apo-10'-carotenal, a process which is also not observed.

6.6 Inhibition of *A. thaliana* CCD8

As discussed, phenotypic effects on lateral shoot branching have been observed in *A. thaliana* on application of the hydroxamic acid compounds D2, D4, D5 and D6, whereby application of inhibitor results in an increase in the number of lateral shoots, indicating inhibition of the biosynthesis of the shoot branching hormone strigolactone (Sergeant *et al*. 2009).

AtCCD8.2 was initially tested for inhibition against compound D2, using the continuous UV method detailed in Section 6.2. Compound D2 was pre-incubated AtCCD8.2 for varying lengths of time before the enzyme substrate, 9-*cis*-*β*-apo-10'-carotenal was added. Analysis revealed that, just as with the NCED subfamily, AtCCD8.2 exhibits a time dependent inhibitory effect. When pre-incubated for 10 minutes there is 75% inhibition of AtCCD8.2. However, following 30 minutes pre-incubation this increases to 100% inhibition of enzyme activity (Figure 6.20).

Therefore the library of hydroxamic acids was screened at a concentration of 100 µM against AtCCD8.2 with 30 minutes pre-incubation of inhibitor using the UV assay on the plate

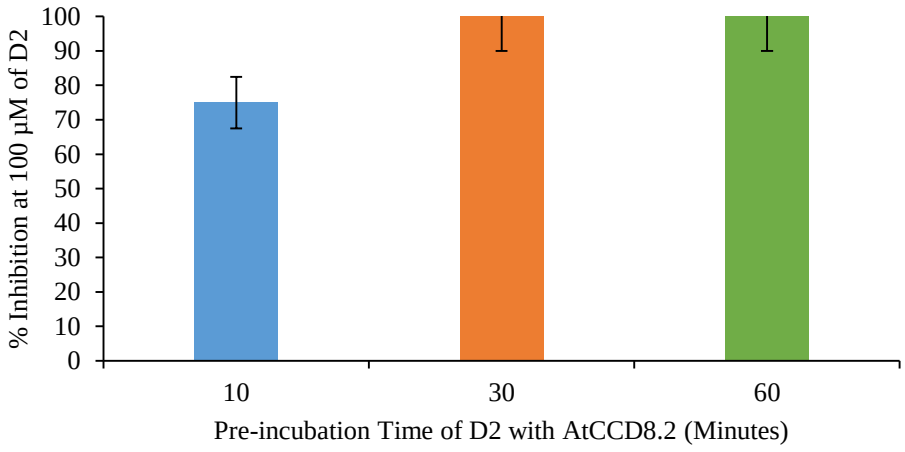


Figure 6.20 – Levels of inhibition observed for compound D2 at 100 µM against AtCCD8.2 with varying lengths of pre-incubation of inhibitor with enzyme.

reader as described in Section 6.2. Results of the assays are detailed in Tables 6.3 to 6.7 and summarised graphically in Figure 6.21. A representative UV assay chromatograph is shown in Figure 6.22.

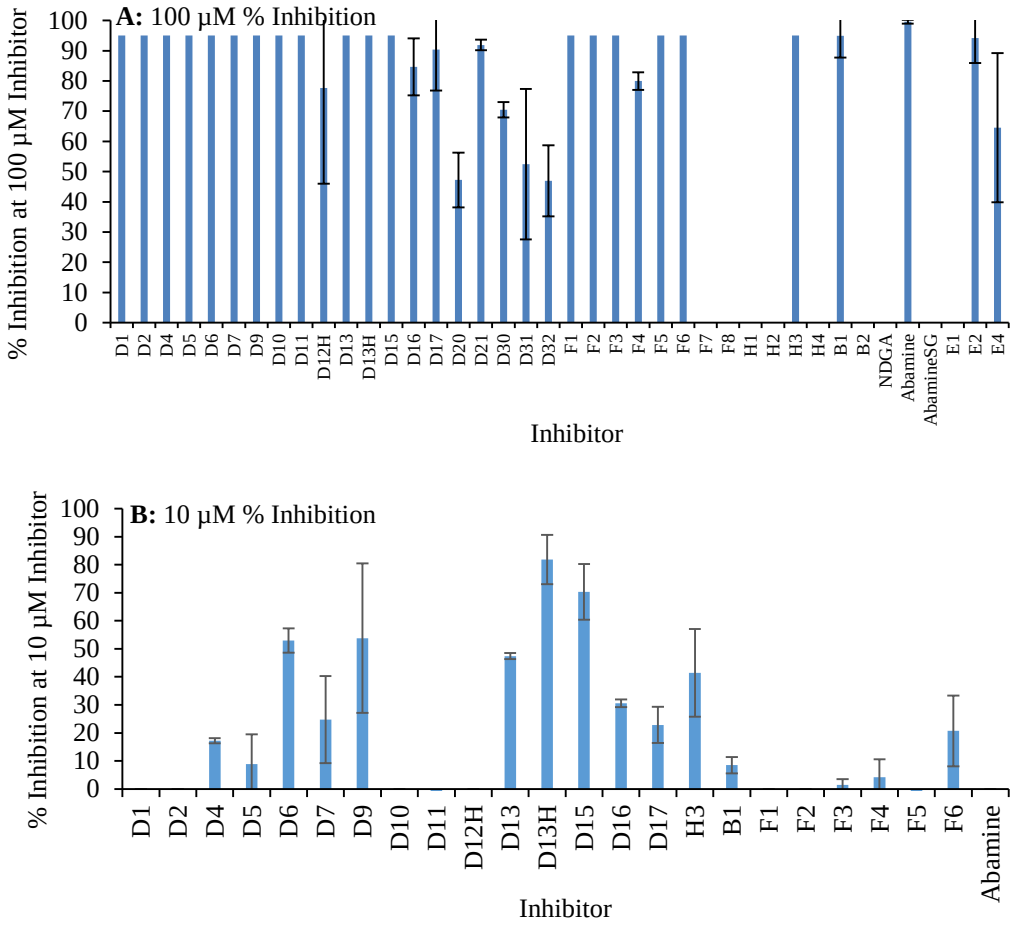


Figure 6.21 – Summary of inhibition data of inhibitors assayed against AtCCD8.2 at 100 µM (**A**, top) and of inhibitors assayed against AtCCD8.2 at 10 µM (**B**, bottom).

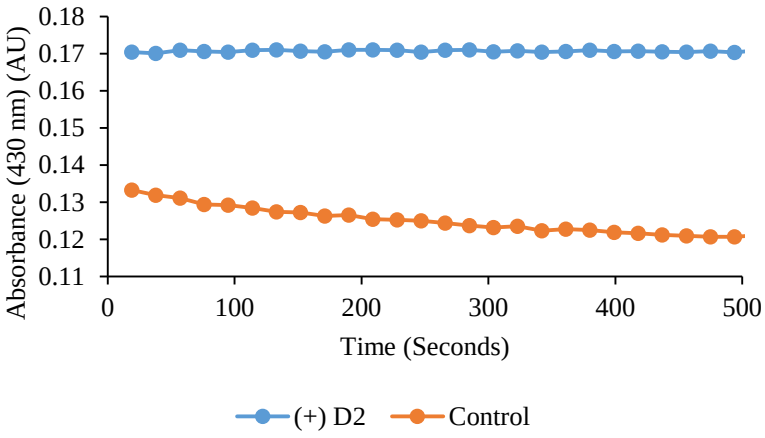


Figure 6.22 – Representative UV chromatograph for the reaction of AtCCD8.2 with 9-*cis*-*β*-apo-10'-carotenal in the absence (control, orange) and presence (blue) of 100 µM D2.

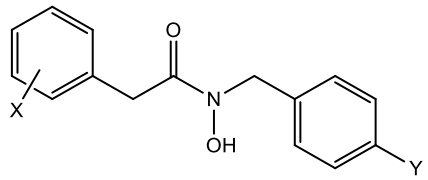
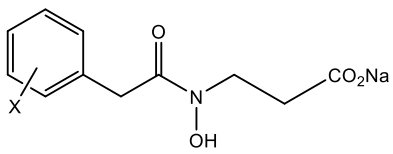
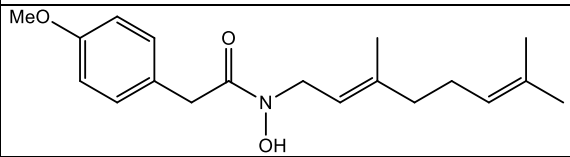
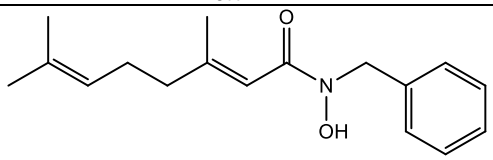
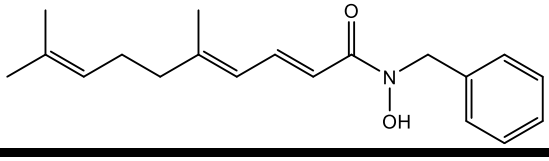
Inhibitor	Structure	Substituent X	Substituent Y	AtCCD8.2 Inhibition at 100 µM (%)	AtCCD8.2 Inhibition at 10 µM (%)
D1		4-Hydroxy	H	>95	0
D2		4-Hydroxy	F	>95	0
D4		4-Methoxy	F	>95	17
D5		3,4-Dimethoxy	H	>95	9
D6		3,4-Dimethoxy	F	>95	53
D7		4-Methoxy	F	>95	25
D16		4-*o*-Butyl	H	84	---
D17		4-Acetyl	H	90	---
D20		4-Methoxy	--	47	---
D21		Napthyl	--	92	---
D30		---	--	70	---
D31		---	--	52	---
D32		---	--	47	---

Table 6.3 – Inhibition data for compounds D1-D7, D20, D21 and D30-D32 versus AtCCD8.2.

At a concentration of 100 µM there is complete inhibition with many of the 'D' series inhibitors. 100% inhibition is seen with compounds D1-D7. Compounds D16 and D17 do not show complete inhibition at 100 µM, which may be due to a steric effect as a result of the

larger acetyl and *O*-butyl substitutions on the phenyl group. None of the inhibitors D1-D7 and D15-D17 showed complete inhibition at a concentration of 10 µM. The carboxylated inhibitors D20 and D21 showed moderate levels of inhibition at 100 µM. The negative charge of the carboxyl may be perturbing interactions within the active site. Fluorinated compounds also showed slightly higher levels of inhibition compared to their unflourinated counterparts. This is slightly unexpected, since *in vitro* fluorination should have little effect. It could be that the fluorine is adjusting the electron density within the phenyl group to aid interactions within the active site, or that the fluorine is acting in a hydrogen bonding capacity. None of the longer compounds, D30-D32, showed complete inhibition at 100 µM (70%, 52% and 47% respectively). This indicates that the active site may be somewhat restricted in size.

In contrast to the compounds in Table 6.3, those in Table 6.4 have an ethylene bridge between the carboxyl of the hydroxamic acid and the aromatic group (compounds in Table

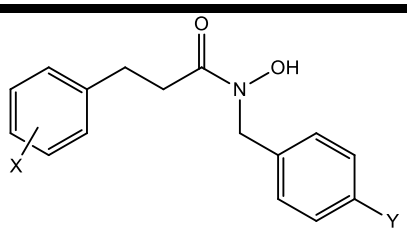
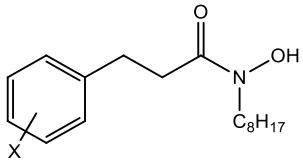
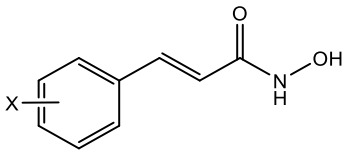
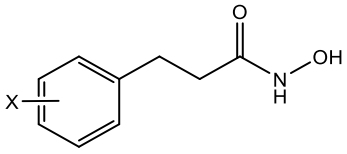
Inhibitor	Structure	Substituent X	Substituent Y	AtCCD8.2 Inhibition at 100 µM (%)	AtCCD8.2 Inhibition at 10 µM (%)
D9		4-Methoxy	H	>95	54
D15		4-Methoxy	F	>95	71
D10		3,4-Dimethoxy	---	>95	0
D11		4-Methoxy	---	>95	0
D12		3,4-Dimethoxy	---	>95	56
D13		4-Methoxy	---	>95	47
D12H		3,4-Dimethoxy	---	78	---
D13H		4-Methoxy	---	>95	82

Table 6.4 – Inhibition data for compounds D9-D13, D12H, D13H and D15 versus AtCCD8.2.

6.3 have a methylene bridge). At 100 μM concentration complete inhibition was seen for compounds D9-D13, D15 and D13H. The strongest levels of inhibition were observed for compounds D13H and D15 (82% and 71% inhibition respectively at 10 μM). The increased chain length provided by the ethylene linker may be positioning the aromatic group for a π-π stacking interaction in the active site. The long alkyl chain on compounds D10 and D11 may perturb inhibition at 10 μM through steric effects. Methoxylated compounds showed higher levels of inhibition than hydroxylated ones, suggesting a possible hydrogen bond donor interaction. As observed previously, fluorinated analogues showed slightly higher levels of inhibition (compounds D15 versus D9).

Compounds D12, D12H, D13 and D13H all have a protonated *N*-alkoxy group, making them slightly smaller. Loss of a substituent on the *N*-alkoxy terminus of the hydroxamic acid does not have a negative effect on the levels of inhibition seen. This could suggest that the inhibition observed is due more to chelation of the active site iron and positioning of the other aromatic group, and less to do interactions of the substituents on the inhibitor to residues within the active site. Compounds D13 and D13H (4-methoxy) both showed greater inhibition than D12 and D12H (3,4-dimethoxy), suggesting a steric effect due to the meta substituted methoxy group. In general, increased inhibition was observed with 4-methoxy substituted inhibitors than with 3,4-dimethoxy substituted inhibitors, supporting the conclusion of a steric clash in the active site.

Of the 'F' series compounds (Table 6.5), 100% inhibition at 100 μM is seen with compounds F1-F6. It can again be seen that fluorinated derivatives with a 3,4-dimethoxy group showed higher levels if inhibition (albeit marginal). When compared to the equivalent 'D' series compounds, the 'F' series compounds showed less inhibition. Although the trend is for small compounds to show greater inhibition, it may be that these 'F' series compounds are too small to bind effectively to the active site. Interestingly, no inhibition at 100 μM is seen with compounds F7 and F8, whilst the highest level of inhibition for the 'F' series was seen

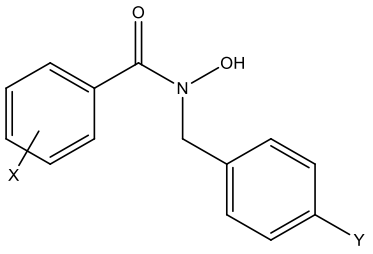
Inhibitor	Structure	Substituent X	Substituent Y	AtCCD8.2 Inhibition at 100 µM (%)	AtCCD8.2 Inhibition at 10 µM (%)
F1		4-Methoxy	H	>95	0
F2		4-Methoxy	F	>95	0
F3		3,4-Dimethoxy	H	>95	1
F4		3,4-Dimethoxy	F	>95	4
F5		3-Chloro	H	>95	0
F6		3-Amino	H	>95	21
F7		3-Bromo	H	0	---
F8		3-Chloro-4-Methoxy	H	0	---

Table 6.5 – Inhibition data for hydroxamic acid compounds F1-F8 versus AtCCD8.2.

with compound F6, a 3-amino substituted compound. The 3-amino group of F6 may be acting as a hydrogen bond donor, enhancing the binding within the active site.

As with compounds D30-D32, the compounds H1-H4 were also poor inhibitors of CCD8. In fact, H1, H2 and H4 showed no inhibition at 100 µM. Compound H3 on the other

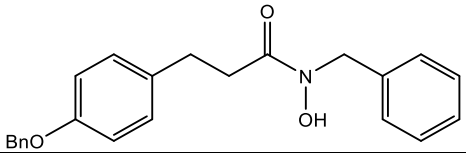
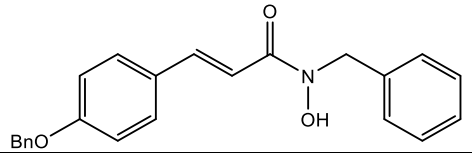
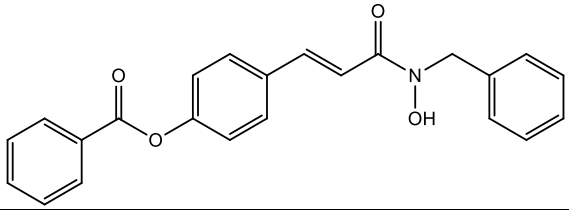
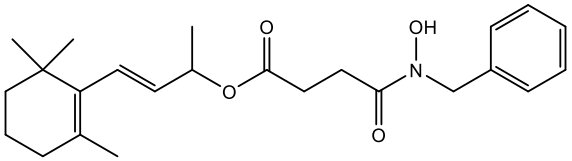
Inhibitor	Structure	AtCCD8.2 Inhibition at 100 µM (%)	AtCCD8.2 Inhibition at 10 µM (%)
H1		0	---
H2		0	---
H3		>95	41
H4		0	---

Table 6.6 – Inhibition data for hydroxamic acid compounds H1-H4 versus AtCCD8.2. Bn represents a benzyl group (CH_2Ph).

hand showed 41% inhibition at 10 μM, which is unusual for such a large compound. It is not immediately clear why H3 is so much more potent with respect to the other members of the H series, the only major difference being the addition of an acetyl group, which may be providing a hydrogen bonding interaction in the active site. It may be though that hydrolysis of the ester in H3 is occurring. This would in effect form a compound similar to D12-D13, which are reasonably potent inhibitors of CCD8.

The hydroxamic acid inhibitor B2 showed no inhibition at 100 μM, further confirming the observation that larger compounds are weaker inhibitors of AtCCD8.2. Compound B1, a smaller compound, on the other hand showed 94% inhibition at 100 μM. The conjugated triene in B1 is likely performing an analogues π-π stacking interaction to the aromatic group in other compounds.

Several non-hydroxamic acid based inhibitors were also assayed against AtCCD8.2. Abamine showed 100% inhibition of AtCCD8.2 at 100 μM, whilst the derivative abamineSG showed no inhibition. The increased chain length on the ester moiety appears to diminish the inhibitory effects, again supporting the observation that larger compounds are weaker inhibitors. Compounds E1, E2 and E4 are tertiary amines like abamine and abamineSG. None of these compounds showed complete inhibition of AtCCD8.2 at 100 μM. When compared to abamine, the loss of the double bond in the structure causes a decrease in the level of inhibition, possibly due to weaker π-π interactions within the active site. Finally, the compound NDGA showed no inhibition at 100 μM. NDGA is reasonably large, which appears to be detrimental to inhibition. It is possible that NDGA could have coordinated to the active site iron through either of the two catechol groups. However, this is apparently not observed, suggesting that either a hydroxamic acid moiety or tertiary amine is required for inhibition.

It is suspected that the time dependent inhibitory effect exhibited by CCD8 towards the hydroxamic acids is a result of the same reasons detailed for NCED; longer incubation of the enzyme with inhibitor allows stronger binding of the inhibitor to the enzyme active site.

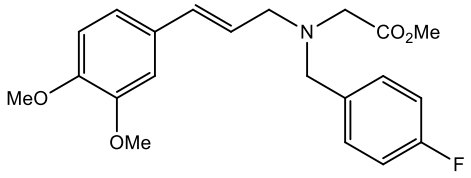
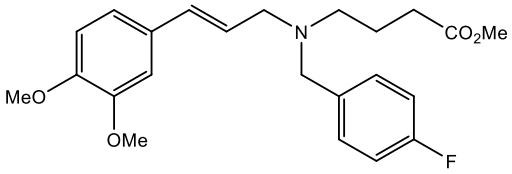
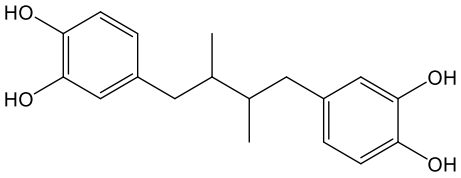
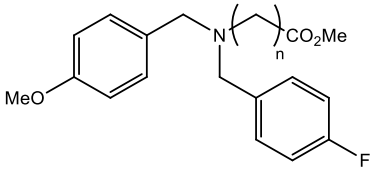
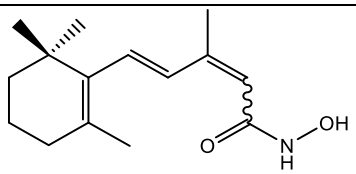
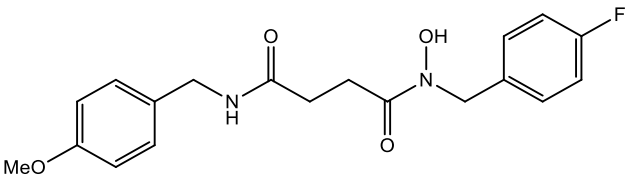
Inhibitor	Structure	Value 'n'	AtCCD8.2 Inhibition at 100 µM (%)	AtCCD8.2 Inhibition at 10 µM (%)
Abamine		--	>95	0
AbamineSG		--	0	---
NDGA		--	0	---
E1		1	0	---
E2		2	94	---
E4		4	65	---
B1		--	94	---
B2		--	0	---

Table 6.7 – Inhibition data for abamine, abamineSG, NDGA, B1, B2 and E1-4 versus AtCCD8.2.

Based on the inhibition observed *in vitro*, several conclusions can be made about the features of hydroxamic acids which are required for inhibition. In general, smaller compounds are better inhibitors of AtCCD8.2 than larger compounds, up to a point. This suggests that the active site of AtCCD8.2 may be quite small, preventing larger compounds from binding. Compounds with an ethylene bridge on the carboxy terminus of the hydroxamic acid were more effective than those with no linker or a methylene linker. Given that the CCD8 cleavage reaction is occurring further along the polyene chain of the apocarotenoid (13,14 for the first

cleavage, compared to a 9,10 cleavage for CCD1/CCD7 and 11,12 for NCED), then the increased alkyl chain length in the hydroxamic acids maybe mimicking this effect, effectively positioning the phenyl group for a π-π stacking interaction.

Smaller groups on the *N*-alkoxy terminus of the hydroxamic acid lend to greater levels of inhibition. Larger groups like carboxylates, geranyl substituents and long alkyl chains were not well tolerated, whilst smaller groups such as phenyl groups and protons were. Fluorinated derivatives are also more effective inhibitors, although it is not clear why. It may be that the fluorine substituent is hydrogen bonding to residues within the active site. On the carboxyl terminus of the hydroxamic acid a 4-methoxy substitution was preferred. Hydroxyl groups and 3,4-dimethoxylation was less well tolerated, suggesting a possible steric interaction within the active site and possibly a hydrogen bonding interaction. The hydroxamic acids were more potent inhibitors than the tertiary amines.

It is also interesting to note that compound D15 showed 71% inhibition at a concentration of 10 μM, given that D15 has also been shown to exhibit a lateral root branching phenotype in *A. thaliana* (addition of D15 causes a decrease in the number of lateral roots) (van Norman *et al*. 2013). The biochemical basis of this phenotype is unknown and was proposed to be due to an unknown carotenoid derived signalling molecule. However, given that D15 has shown activity against CCD8, it is possible that the phenotype could be a strigolactone mediated process, or due to a signalling molecule derived from 9-*cis*-*β*-apo-10'-carotenal, which would accumulate *in planta* on inhibition of CCD8. Unfortunately, due to limitations with regards to the amount of compound available, it has not been possible to test any of the inhibitors in Table 6.4 *in planta*.

6.7 Conclusions

Although it has not been possible to definitively identify the nature of the reaction products, AtCCD8 does appear to catalyse the selective cleavage of 9-*cis*-*β*-apo-10'-carotenal. AtCCD8.2 exhibits substrate inhibition at high concentrations, an effect not observed with any

other member of the CCD enzyme family. Such an effect may make labelling experiments and the detection of products and intermediates difficult. The mechanism of CCD8 cleavage of 9-*cis*-*β*-apo-10'-carotenal cleavage is still an issue of debate. However, AtCCD8 has been shown here to proceed through a two-step reaction, requiring acid-base catalysis and the participation of at least different active site residues; a histidine, aspartic acid or glutamic acid and a cysteine. Based on this data a possible mechanism for CCD8 has been proposed involving a redox active cysteine residue.

Inhibition has been observed of the CCD8 reaction in the presence of the hydroxamic inhibitors. *In planta*, phenotypic effects on lateral shoot branching were observed with compounds D2, D4, D5 and D6. *In vitro*, 100% inhibition of CCD8 is observed against these inhibitors at 100 µM. Some inhibition is also seen at 10 µM. This data would suggests that the *in planta* phenotypic effects on lateral shoot branching are due to inhibition of CCD8. Other members of the hydroxamic acid inhibitor family have also been shown to inhibit AtCCD8 and it has been shown that smaller compounds in general are more effective inhibitors of AtCCD8. This could imply that the AtCCD8 active site is quite small, which could have implications for the mechanism of CCD8 (Figure 6.23).

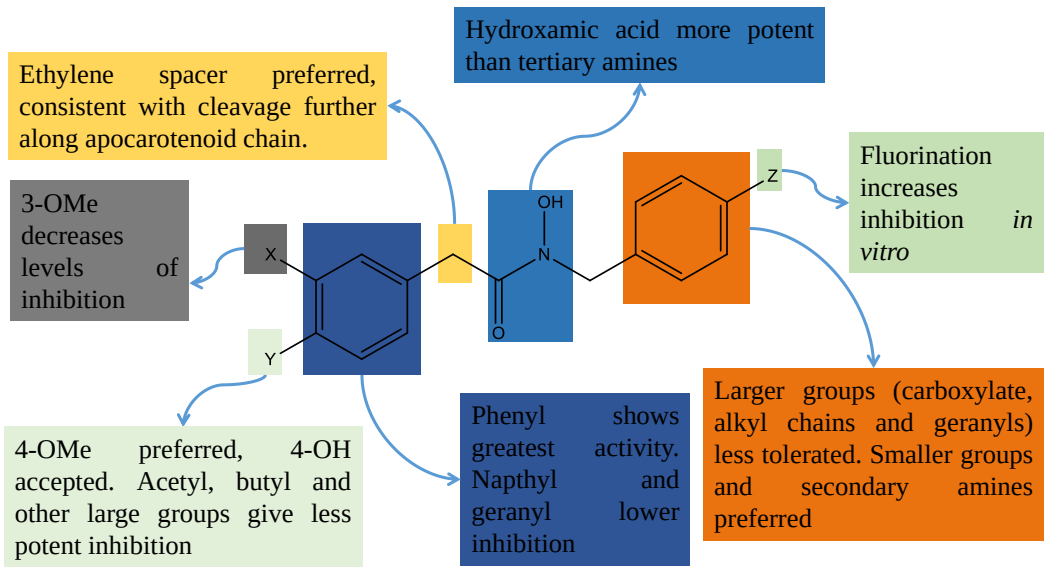


Figure 6.23 – Summary of structure-activity relationship data observed for the hydroxamic acid inhibitors against AtCCD8.2.

6.8 Future Work

As stated, the mechanism by which CCD8 produces carlactone is still a matter of some debate. In order to provide further insight it would be necessary to obtain a crystal structure of CCD8, ideally with a bound ligand, which would allow the identification of residues which are potentially involved in catalysis. A crystal structure would also help to validate results obtained from incubations with group specific reagents. It would also be necessary to identify intermediates by mass spectrometry in order to elucidate the CCD8 mechanism. This can be performed via rapid quench experiments. If CCD8 does proceed through a radical mechanism involving a redox active cysteine, then it may be possible to inhibit the CCD8 reaction via incubation with a radical, such as (2,2,6,6-Tetramethyl-piperidin-1-yl)oxyl (TEMPO). Radical traps could also be used to probe the mechanism of CCD8.

Further work could also be performed with regards to developing a selective inhibitor of shoot branching. The compounds D2, D4, D5 and D6 are active *in vitro* and *in planta*, but have cross reactivity to other members of the CCD enzyme family. Ideally it would be advantageous to have an inhibitor which is selective solely for the CCD8 enzyme family. It would also be of interest to test further hydroxamic acid inhibitors *in planta* to investigate whether they exhibit either the shoot branching phenotype or lateral root branching phenotype.

Chapter Seven: Conclusions and Future Work

7.1 Conclusions

A summary of results is presented in Table 7.1. In their 2009 paper, Sergeant *et al*. reported phenotypic effects on shoot branching upon the application of certain hydroxamic acid inhibitors to *Arabidopsis thaliana* (Sergeant *et al*. 2009). One such phenotypic effect observed was an effect on seed germination, whereby in the presence of the hydroxamic acids D2 and D4 in *A. thaliana* there was a decrease in the time taken for seeds to germinate. This effect was believed to be mediated through the inhibition of the enzyme NCED, which is involved in the biosynthesis of the plant hormone abscisic acid, which is a promoter of

Enzyme	Conclusions	Most effective inhibitors
LeCCD1	CCD1 can be purified by removal of the active site iron. Structure-activity analysis has identified key features of hydroxamic acid inhibitors required for inhibition.	D1 (IC50 = 0.9 µM) D2 (IC50 = 0.8 µM) D3 (IC50 = 0.8 µM)
LeNCED1	Tomato NCED1 is inhibited weakly by the hydroxamic acids. Structure-activity analysis has identified key features of hydroxamic acid inhibitors required for inhibition.	D30 (42% at 100 µM) D8 (40% at 100 µM) D4 (33% at 100 µM)
ZmNCED1, 2, 3A, 3B & 9	There is no significant difference between the levels of inhibition observed by the *Z. mays* NCED homologs verses hydroxamic acids inhibitors, hence it is not possible to identify which homolog is the cause of the seed germination phenotype. ZmNCED1 exhibits a time dependent inhibition effect towards the hydroxamic acids however.	100% inhibition of ZmNCED1 vs. D2 and D4 at 100 µM with pre-incubation
AtABA2	ABA2 was cloned and expressed *in vitro*. However, it was not possible to develop a high throughput coupled NCED1-ABA2 assay for the screening of inhibitors.	N/A
AtCCD7	CCD7 is not inhibited by the hydroxamic acid compounds *in vitro* and hence is not the cause of the shoot branching phenotype observed *in planta.*	No inhibition observed
OsD27	D27 appears to be a [2Fe2S] protein which isomerises β-carotene through a radical mediated mechanism. D27 shows weak inhibition against the hydroxamic acids and hence is not the cause of the shoot branching phenotype seen in planta.	D30 (41% at 100 µM) D12H (40% at 100 µM) B2 (40% at 100 µM)
AtCCD8	AtCCD8 is inhibited by hydroxamic acids at concentrations as low as 10 µM and is inhibited by the compounds D2, D4, D5 and D6, which exhibit a phenotypic effect on seed germination *in planta*, indicating that inhibition of CCD8 is the biochemical basis of this phenotype. CCD8 catalyses a two step reaction involving acid base catalysis and a possible redox active cysteine.	D13H (82% at 10 µM) D15 (71% at 10 µM) D12 (56% at 10 µM)

Table 7.1 – Summary of results.

secondary seed dormancy. Despite *in vitro* data not initially indicating that the target of D2 and D4 was NCED, through the data illustrated here it can in fact be concluded that it is likely that the biochemical basis of the seed germination phenotype is indeed due to inhibition of NCED (Figure 7.1). *In vitro* it was observed that NCED exhibits a time dependent inhibitory effect with the hydroxamic acids, whereby pre-incubation of the hydroxamic acid with the enzyme is required for inhibition. This is likely to mimic the effect *in planta* and could well be due to oxidation of the active site iron.

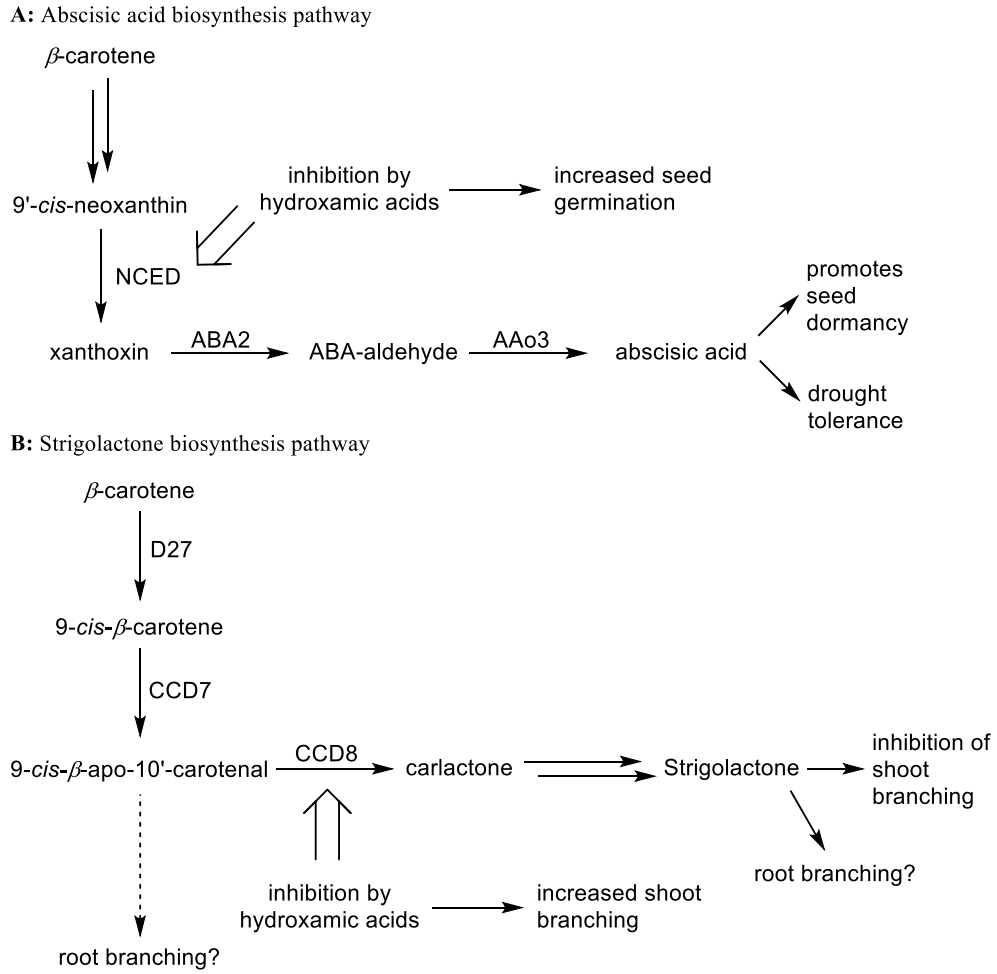


Figure 7.1 – Schematic of the effects mediated by abscisic acid (**A**) and strigolactone (**B**) and the points of intervention by the hydroxamic acids.

Attempts were made to develop a high throughout screen for the identification of new inhibitors of NCED. To this end, the enzyme following NCED on the abscisic acid biosynthesis pathway, ABA2, was cloned and expressed *in vitro*. It was hoped that since ABA2 is an NAD dependent enzyme, the absorbance change at 340 nm could be observed

when ABA2 was coupled to NCED. Unfortunately, an unacceptably high background reaction of unknown origin prevented development of a coupled high throughput screen.

It was also reported by Sergeant *et al*. that the hydroxamic acids D2, D4, D5 and D6 had an effect on lateral shoot branching *in planta*, although the biochemical basis for this phenotype was unknown (Sergeant *et al*. 2009). In order to investigate this the first three enzymes on the strigolactone biosynthesis pathway, D27, CCD7 and CCD8, were expressed and assayed *in vitro*. It was observed that there was no inhibition of CCD7 *in vitro* by the hydroxamic acids (Figure 7.1). Some very low level inhibition of D27 was observed for a handful of compounds but CCD8 showed 100% inhibition at 100 µM in the presence of several of the hydroxamic acids. Additionally, CCD8 was inhibited *in vitro* by the compounds D2, D4, D5 and D6 which showed a phenotype *in planta*. Some compounds showed inhibition *in vitro*, but did not show an effect *in planta* (D1, D3 and D7). This is likely to be due to issues associated with uptake of molecules into the plant, transport within the plant and metabolism of the inhibitors within the plant. Hence it can be concluded that the biochemical basis of the shoot branching phenotype is due primarily to inhibition of CCD8. Compound D15 showed 71% inhibition of AtCCD8 at 10 µM *in vitro* and has also been shown to exhibit a phenotypic effect of lateral root branching in *A. thaliana*. It is possible that CCD8 is involved in the biosynthesis of a signalling molecule which controls this process.

It is possible, both in the case of the shoot branching phenotype and the seed germination phenotype, that the hydroxamic acids are inhibiting other enzymes. However, with regards to the non-CCD enzymes investigated during this project, ABA2 and D27, little to no inhibition was seen in the presence of the hydroxamic acid inhibitors.

A library of hydroxamic acid inhibitors was screened against the CCD enzymes CCD1, NCED1, CCD7 and CCD8 and it was possible to identify the key features of the hydroxamic acids required for the inhibition of these enzymes. CCD1 showed the highest levels of inhibition and it is interesting to note that CCD1 shows the highest substrate

promiscuity out of all the CCD enzymes. If it has a larger active site that may explain why it is most amenable to inhibition by the hydroxamic acids. Time dependent inhibition has been observed with regards to the hydroxamic acids towards two members of the CCD enzyme family: CCD8 and NCED. This effect may be due to oxidation of iron (II) in the enzyme to iron (III), which binds hydroxamic acids more strongly. This effect has also been seen with *p*-hydroxy-phenylpyruvate dioxygenase from *A. thaliana* (Sergeant *et al*. 2013).

Several observations have also been made with regards to the mechanism of CCD enzymes, in particular CCD8. Anecdotally, with the CCD7 cleavage it appears that a degree of isomerisation occurs during the cleavage reaction, with a number of other smaller peaks evident following separation of the products by HPLC, although these peaks could not be definitely identified as isomers of the CCD7 cleavage product. In this regard CCD7 is quite similar to the CCD enzyme NinaB (Oberhauser *et al*. 2008), which also catalyses an isomerisation of retinal during the cleavage of *β*-carotene. With regards to the CCD8 cleavage mechanism, although it has not been possible to definitively prove the mechanism of carlactone formation, it has been observed that the CCD8 mechanism does utilise acid-base catalysis in some form. Moreover, stopped flow analysis has shown that CCD8 uses a two-step mechanism and at least three different active site residues are required for catalysis. Using the data obtained a possible mechanism for CCD8 has been postulated involving a redox active cysteine.

The enzyme D27 was identified as a *β*-carotene isomerase by Alder and co-workers (Alder *et al*. 2012), but since its discovery there have been no biochemical investigations of the enzyme. Based on observations made here we hypothesise that D27 is an iron-sulfur protein, which isomerises all-*trans*-*β*-carotene to 9-*cis*-*β*-carotene through one electron transfer either to or from the *β*-carotene substrate. D27 also appears to exhibit product inhibition, which suggest that *in planta*, D27 is closely associated with the second enzyme on the strigolactone biosynthesis pathway, CCD7.

7.2 Future Work

Based on the structure activity relationships obtained, it would be of interest to develop a selective inhibitor for each member of the CCD enzyme family. This would be of use in an agrochemical setting as ideally only one enzyme would be inhibited in obtaining a given phenotype. It may also be of interest to explore the other enzymes on the abscisic acid and strigolactone biosynthesis pathways to determine if they are effected by the hydroxamic acid inhibitors, or whether other CCD family members such as CCD4 are inhibited. The hydroxamic acid inhibitors developed could also be used to explore the roles of CCDs in other organisms in a chemical genetics approach. Applying an inhibitor and observing whether phenotype is present could help to identify functions of carotenoid derived molecules.

Despite the progress made here, there is much work to be done to enable a detailed understanding of the biochemistry of the CCD enzyme family. An important step could be obtaining a crystal structure of a CCD enzyme with either bound substrate or inhibitor. This would be especially useful with regards to elucidation of the CCD8 mechanism. Further work to understand the mechanism CCD8 should focus around rapid quench experiments and determining the nature of intermediates produced. Finally, more work is required to fully understand the mechanism by which D27 works. Experiments such as cyclic voltammetry should reveal whether a redox process is occurring during the mechanism whilst advanced spectroscopic methods such as electron paramagnetic resonance and Mossbauer spectroscopy could be used to determine the nature of the iron containing active site.

Chapter Eight: Experimental

8.1 General Experimental Information

All reagents and solvents were, unless otherwise stated, used as received and purchased from VWR, Sigma-Aldrich UK, Fischer Scientific, MP Biomedical or Invitrogen. All procedures were conducted at room temperature (≈ 23° C) unless specifically stated, all biological procedures were conducted under sterile conditions and all solvents used for HPLC were of HPLC grade and degassed prior to use.

Arabidopsis thaliana CCD7 and CCD8, *Solanum lycopersicum* CCD1a and NCED1 and *Zea mays* NCED1, NCED2, NCED3A, NCED3B and NCED9 were provided by Dr Andrew Thompson and Jake Chandler (Cranfield University). A summary of the enzymes and accession numbers is shown in Table 8.1.

Authentic 9'-*cis*-neoxanthin was obtained by Syngenta from CaroteNature (Ostermundigen, Switzerland) and compounds D2, D4, abamine and abamineSG were

Enzyme	Working BL21 Construct Used	Accession Number for Enzyme
LeCCD1a	pGEX-4T-1-LeCCD1a pRosetta BL21	Q6E4P5
LeNCED1	pGEX-4T-1-LeNCED1 pRosetta BL21	O24023
ZmNCED1	pGEX-4T-1-ZmNCED1 BL21	O24592
ZmNCED2	pGEX-4T-1-ZmNCED2 BL21	N/A
ZmNCED3A	pGEX-4T-1-ZmNCED3A BL21	N/A
ZmNCED3B	pGEX-4T-1-ZmNCED3B BL21	N/A
ZmNCED9	pGEX-4T-1-ZmNCED9 BL21	N/A
AtABA2	pET-200-AtABA2 BL21	Q9C826
OsD27	pGEX-4T-1-OsD27.1 BL21	C7AU21
AtCCD7	pGEX-4T-1-AtCCD7.2 pRosetta pBB541 BL21	Q7XJM2
AtCCD8	pET-151-AtCCD8.3 pRosetta BL21	Q8VY26

Table 8.1 – Working bacterial constructs for enzymes used. Accession numbers for ZmNCED2, 3A, 3B and 9 are not applicable as sequences have not been uploaded to a database.

provided by David Brocklehurst (Syngenta, Jealott's Hill).

Hydroxamic acids were used from liquid stocks checked via mass spectrometry analysis for authenticity before use. Compounds were synthesised and screened against LeCCD1a and LeNCED1 by the following: compound A1 was synthesised and assayed by Jian-Jun Li. Compound B1 was synthesised and assayed by Christine Fox. Compound B2 was synthesised and assayed by Peter Harrison. Compounds E1-4 were synthesised by Mohsina Khan and assayed by Peter Harrison. Compounds D1-D7 were synthesised and assayed by Jian-Jun Li. Compounds D8-D11 were synthesised and assayed by Christine Fox. Compounds D14-15 were synthesised and assayed by Peter Harrison. Compounds D16-D17 were synthesised by Samuel Lowe and assayed by Peter Harrison. Compounds D12-D13 were synthesised and assayed by Christine Fox. Compounds D12H and D13H were synthesised by Paul Sainsbury. Compounds D20-D21 were synthesised and assayed by Peter Harrison. Compounds D30-D32 were synthesised and assayed by Sarah Shepherd. Compounds H1-H4 were synthesised by Jian-Jun Li and assayed by Peter Harrison Compounds F1-F4 were synthesised and assayed by Nicola Brookbank. Compounds F5-F6 were synthesised by Robert Jenkins and assayed by Peter Harrison. Compound F7-F8 were synthesised by Nathaniel Hsueh and assayed by Peter Harrison.

8.2 Instruments and Equipment

Sterilisation of media and equipment for the growth of bacterial strains was performed in a Prioclave autoclave according to standard procedures.

Centrifugation was performed in a Sorval Rc 6 Plus centrifuge with an SLA-3000 or SS-34 rotor (for large volumes), an Eppendorf 5810R or 5804R (for 15-50 mL volumes) and either a Thermo AccuSpin micro or an Eppendorf Centrifuge 5415R (for samples less than 2 mL).

PCR was performed using an Eppendorf Mastercycler Personal. Agarose gel electrophoresis was performed using a BioRad Mini-Sub Cell GT system and visualised at

260 nm using a UVP Ultraviolet (UV) transilluminator. Fermentas FastRuler low range ladder was used to estimate deoxyribonucleic acid (DNA) fragment sizes.

Overnight cultures were grown in either an Innova 44 or an Innova 4300 shaker. The optical density (OD) at 600 nm was recorded with a BioMate3 ultraviolet-visible (UV-Vis) spectrometer. Cell lysis was performed with either a Sonics Vibra Cell sonicator at 50 Hz or a Constant Systems pneumatic cell disruptor at 20.1 kilo-pound per square inch (kpsi). A GE Healthcare AKTA design fast protein liquid chromatograph (FPLC) was used for FPLC purification equipped with a P-920 pump and UPC-900 UV monitor. Sodium-dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed using a BioRad Mini Protean II electrophoresis system and visualised using bromophenol blue G-250. Thermo Scientific PageRuler Plus prestained protein ladder was used for all SDS-PAGE gels (unless stated otherwise).

High-performance liquid chromatography (HPLC) was performed with either an Agilent 1100 Series or 1200 Series liquid chromatograph (Agilent Technologies, Santa Clara, USA) equipped with a G1311A quaternary pump (both 1100 and 1200) and a G1315A photodiode array detector (1100) or a G1314B variable wavelength detector (1200). Liquid-chromatography mass spectrometry (LC-MS) was performed with an Agilent 1200 HPLC system equipped with a G1315A photodiode array detector coupled to a Bruker HTC-Ultra electrospray ionisation (ESI) mass spectrometer. Data analysis was performed with either Bruker Data Analysis post-processing software or Agilent ChemStation (B.01.03).

Low resolution ESI was performed using an Agilent 6130B single Quad mass spectrometer. High resolution mass spectra were obtained via the departmental Mass Spectrometry Service using a Bruker MaXis mass spectrometer or Bruker MicroTOF mass spectrometer.

Plate reader assays were performed with either a Tecan (Männedorf, Switzerland) GENios Microplate reader (Tecan) or a Tecan infinite 200Pro using Grenier or NUNC transparent polystyrene 96 well plates (for UV), PerkinElmer polystyrene white 96 well plates (for luminescence) and Grenier black polystyrol 96 well plates (for fluorescence). UV-Vis measurements were obtained using either a Varian Cary 1 or Varian Cary50 Bio UV-Vis spectrometer.

8.3 General Buffers and Solutions

Buffers and procedures are adapted from Sambrook & Russell, 2001.

8.3.1 Tris-Borate EDTA (TBE) Buffer

70 mM Tris-base, 90 mM boric acid, 8 mM ethylenediaminetetraacetic acid, pH 8.0.

8.3.2 Luria Broth (LB)

1% yeast extract (w/v), 1% tryptone, (w/v), 1% sodium chloride (w/v). For LB agar: additional 2% agar (w/v).

8.3.3 SDS-PAGE Loading Dye

60 mM Tris-HCl, 10% glycerol (v/v), 2% sodium dodecyl sulfate (w/v), 0.1% bromophenol blue (w/v), 1% 2-mercaptoethanol (v/v).

8.3.4 SDS-PAGE running buffer

150 mM glycine, 20 mM tris base, 0.8% sodium dodecyl sulfate (w/v).

8.3.5 SDS-PAGE Staining Solution

50% distilled water (v/v), 40% methanol (v/v), 10% acetic acid (v/v), 1% Coomassie R250 (w/v).

8.3.6 SDS-PAGE Destaining Solution

70% distilled water (v/v), 20% methanol (v/v), 10% acetic acid (v/v).

8.3.7 Tris Buffer pH 8.8

1.5 M Tris-HCl, pH 8.8.

8.3.8 Tris Buffer pH 6.8

500 mM Tris-HCl, pH 6.8.

8.3.9 Western Blotting Wet Transfer Buffer

25 mM Tris-HCl, 192 mM glycine, 20% methanol (v/v).

8.3.10 PBST Buffer

PBS buffer plus 0.1% Tween-20 (v/v).

8.3.11 Phosphate Buffered Saline (PBS)

140 mM sodium chloride, 2.7 mM potassium chloride, 10 mM disodium hydrogen phosphate, 1.8 mM potassium dihydrogen phosphate, pH 7.3.

8.3.12 GST Elution Buffer

50 mM Tris-HCl, 20 mM reduced glutathione, pH 8.0.

8.3.13 N-His_6 Wash Buffer

50 mM sodium hydrogen phosphate, 300 mM sodium chloride, 20 mM imidazole, pH 8.0.

8.3.14 N-His_6 Elution Buffer

50 mM sodium hydrogen phosphate, 300 mM sodium chloride, 250 mM imidazole, pH 8.0.

8.3.15 MBP Lysis Buffer

PBS buffer plus 25 mM DTT, lysozyme (0.5 mg mL^{-1}), pH 7.4.

8.3.16 MBP Wash Buffer

20 mM Tris-HCl, 200 mM sodium chloride, 1 mM EDTA, 10 mM BME, pH 7.4.

8.3.17 MBP Elution Buffer

20 mM Tris-HCl, 200 mM sodium chloride, 1 mM EDTA, 10 mM BME, 10 mM maltose hydrate, pH 7.4.

8.3.18 Bis-Tris Buffer

100 mM bis-tris, 10% glycerol (v/v), 0.05% Triton X-100 (v/v), pH 6.7.

8.3.19 MOPS Buffer

20 mM MOPS, 80 mM sodium chloride, pH 7.5.

8.3.20 Potassium Phosphate Buffer

100 mM potassium phosphate, pH 7.2

8.3.21 NCED Assay Buffer

PBS buffer plus catalase (1 mg mL^{-1}), 0.05% Triton X-100 (v/v), 20 mM iron (II) sulfate, 20 mM sodium ascorbate.

8.3.22 CCD7 Assay Buffer

100 mM HEPES, 0.2 mM iron (II) sulfate, 1 mM tris(2-carboxyethyl)phosphine, catalase (1 mg mL^{-1}), 0.05% Triton X-100 (v/v), pH 7.8.

8.3.23 CCD8 Assay Buffer

100 mM HEPES, 1 mM tris(2-carboxyethyl)phosphine, pH 7.8.

8.3.24 Stopped-Flow Assay Buffer

100 mM HEPES, pH 7.8.

8.4 General Procedures

8.4.1 Preparation of Chemically Competent E. coli BL21 and E. coli DH5α

The method used was adapted from the method of Inoue (Inoue *et al*. 1990). *E. coli* DH5α and BL21 were obtained either from Life Technologies or New England Biolabs and plated onto Luria Broth (LB) agar and grown overnight at 37º C, 180 rpm. A single colony was picked and used to inoculate 200 mL of LB media. *E. coli* was grown at 20º C until OD at 600 nm reached 0.6 AU. Cells were harvested at 4220 x g at 4º C for 20 minutes, then washed with ice cold 100 mM calcium chloride (2 x 10 mL) with gentle pipetting before being incubated on ice for 15 minutes. Cells were pelleted by centrifugation as above and resuspended in ice cold 100 mM calcium chloride (1 mL) and left on ice for 30 minutes.

Glycerol (50% v/v) was added to a final concentration of 15% (v/v) and cell aliquots (100 µL) were snap frozen in liquid nitrogen before being stored at -80º C.

8.4.2 Transformation of Chemically Competent E. coli BL21 and E. coli DH5α

A single aliquot (100 µL) of chemically competent BL21 or DH5α cells was defrosted on ice. Plasmid DNA (3 µL, approximately 50–100 ng) was added and mixed by gentle pipetting. Cells were incubated on ice for 60 minutes before being heat shocked at 42º C for 40 seconds and immediately placed on ice before room temperature LB media was added (250 µL). Cells were incubated at 37º C, 180 rpm for 60 minutes and *E. coli* (2 x 150 µL) was then plated onto LB agar containing appropriate antibiotic (see Table 8.2) before being incubated at overnight at 37º C.

Antibiotic	Final Concentration (µg mL^{-1})	Plasmid(s) Used With
Ampicillin	100	pGEX-4T-1 pET-151 pMAL-c5e
Chloramphenicol	35	pRosetta pAC
Kanamycin	35	pET-200 pUC-57-KAN
Spectinomycin	100	pBB541

Table 8.2 – Concentrations of antibiotics required for selection of plasmids.

8.4.3 Overproduction of Proteins in E. coli BL21

LB media (10 mL) containing the appropriate antibiotic (see Table 8.2) was inoculated with a single colony of *E. coli* BL21 (containing desired plasmid(s)) and grown overnight at 37º C, 180 rpm. LB media (500 mL) containing appropriate antibiotics was inoculated with the overnight culture and grown at 37º C, 180 rpm until the OD at 600 nm reached 0.6 AU. Isopropyl *β*-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Cultures were incubated overnight at 20º C, 180 rpm. Cells were harvested at 4220 x g at 4º C for 15 minutes and either frozen at -20º C or used as described in the following chapter.

8.4.4 Agarose Gel Electrophoresis

Agarose (1 g) was dissolved in TBE buffer (100 mL) via heating in a microwave for 2 minutes. Upon cooling, GelRed (Cambridge Bioscience) (3 µL) was added and the molten agarose was added to an appropriate cast with comb. Gels were run at 100 V for 45 - 60 minutes. DNA was visualised at 260 nm.

8.4.5 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Protein samples were analysed using either 8% or 15% SDS-PAGE gels unless otherwise stated. To protein samples (5 µL) was added water (15 µl) and SDS-PAGE loading dye (5 µL). Samples were boiled for 10 minutes at 100° C before centrifugation at 20300 x g for 5 minutes. Gels were made using the recipe detailed in Table 8.3 and run at 200 V for one hour at room temperature. Gels were stained for one hour using SDS-PAGE staining solution and destained for 30 minutes with SDS-PAGE destaining solution.

Gel Percentage	Water (mL)	30% Acrylamide Mix (mL)	Tris Buffer pH8.8 (mL)	Tris Buffer pH 6.4 (mL)	10% (w/v) Sodium dodecyl sulfate (SDS) (mL)	10% (w/v) ammonium persulfate (APS) (mL)	Tetramethylethyl-enediamine (TEMED) (mL)
8%	2.3	1.3	1.3	N/A	0.05	0.05	0.003
15%	1.1	2.5	1.3	N/A	0.05	0.05	0.002
Stacking Gel	0.68	0.17	N/A	0.13	0.01	0.01	0.001

Table 8.3 – Recipes for SDS-PAGE gels used.

8.4.6 PCR Screening of Transformants

To a single colony of transformed *E. coli* was added distilled sterile water (3 µL). PCR was performed according to the procedure detailed for the particular enzyme with appropriate primers. PCR product was analysed via agarose gel electrophoresis. PCR reactions consisted of water (39.1 µL), 10 X *pfx* buffer (Invitrogen) (5.0 µL), 50 mM magnesium sulfate (1.0 µL), 10 mM deoxynucleotide triphosphates (dNTPs) (1.5 µL), 10

μM forward primer (1.5 μL), 10 μM reverse primer (1.5 μL), *pfx* tag polymerase (1.0 μL) and *E. coli* colony solution (1.0 μL).

8.5 *Solanum lycopersicum* CCD1a

8.5.1 Preparation of Cell Free Extract and Purification of LeCCD1a

Cells were prepared as described in Section 8.4.3 using pGEX-4T-LeCCD1a pRosetta *E. coli* BL21 and were resuspended into PBS buffer (10 mL) with 10% glycerol (v/v), 0.05% Triton X-100 (v/v) and 1 mM 2-mercaptoethanol (BME). Lysozyme and dithiothreitol (DTT) were added to a final concentration of 25 μg mL^{-1} and 5 mM respectively. Resuspended cells were incubated at room temperature for 20 minutes before sonication at 50 Hz and cell disruption at 20.1 kpsi. DNAse I was added to a final concentration of 5 μg mL^{-1} and allowed to incubate at room temperature for 20 minutes. The cell free extract was centrifuged at 20300 x g at 4º C for 15 minutes and the supernatant was filtered through a 0.2 μm filter. For assays involving cell free extract, the cell free extract was used without any further purification.

To the filtered cell free extract was added 1, 10 phenanthroline to a final concentration of 1 mM and the cell free extract was incubated on ice for 15 minutes. The solution was added to GST resin (2 mL, GE Life Science) prepared according to the manufacturer's instructions. Cell free extract was passed through the column a total of three times before being washed with PBS buffer (25 mL) with 0.05% Triton X-100 (v/v), 10 % glycerol (v/v) and 1 mM BME. The apo-protein was eluted into GST elution buffer (10 mL) plus 10 % glycerol (v/v), 1 mM BME and 0.05% Triton X-100 (v/v). Apo-protein was then concentrated using a 50 kDa centrifugal filter unit before buffer exchange using a PD10 gel filtration column (GE Life Sciences) used according to the manufacturer's instructions. Apo-protein was eluted into PBS buffer (3.5 mL) plus 10% glycerol (v/v), 1 mM BME and 0.05% Triton X-100 (v/v) and concentrated further in a 10 kDa centrifugal filter. Analysis via SDS-PAGE was performed using an 8% gel. Apo-protein concentration was determined via

the method of Bradford (Bradford 1976). Apo-protein was aliquoted, snap frozen in liquid nitrogen and stored at -80º C.

8.5.2 Assays of LeCCD1a

LeCCD1a assays were conducted in a total volume of 200 µL in a 96 well microtiter plate. The assay substrate solution was prepared by adding 2% (w/v) ethanolic all-*trans*-*β*-apo-8'-carotenal (5 µL) to 4% (w/v) ethanolic *β*-octylglucoside (25 µL). The ethanol was removed under nitrogen to give an orange residue which was resuspended into PBS buffer (150 µL) plus 10% glycerol (v/v) and 10 mM sodium ascorbate. The substrate solution was then allowed to incubate at room temperature for 20 minutes. To the assay solution was added either cell free extract (50 µL) or purified enzyme (20 µL, ≈40 µg) and the reactions were monitored at 485 nm for 60 minutes at room temperature.

For assays involving purified apo-LeCCD1a, the apo-protein was allowed to defrost on ice before iron (II) sulfate and sodium ascorbate were added to a final concentration of 1 mM. The protein sample was then allowed to incubate on ice for a further 10 minutes before being assayed.

Inhibitors were dissolved in either ethanol or dimethyl sulfoxide (DMSO) and a total of 2 µL of inhibitor stock was added to the assay to give a final concentration of between 1 µM to 100 µM.

8.6 *Solanum lycopersicum* and *Zea mays* NCEDs

8.6.1 Preparation of E. coli BL21 Containing ZmNCEDs

E. coli containing the plasmids pGEX-4T-1-ZmNCED1, pGEX-4T-1-ZmNCED2, pGEX-4T-1-ZmNCED3A, pGEX-4T-1-ZmNCED3B and pGEX-4T-1-ZmNCED9 were obtained from Jake Chandler and Andrew Thompson (Cranfield University). Plasmids were extracted from 10 mL overnight cultures using a Thermo Scientific plasmid extraction kit according to the manufacturer's instructions. ZmNCED plasmids were used to transform *E. coli* BL21 according to Section 8.4.2.

8.6.2 Preparation of Cell Free Extract and Purification of LeNCED1 and ZmNCEDs

Cells were prepared as described in Section 8.4.3 with the appropriate *E. coli* strain and were resuspended into PBS buffer (10 mL) with 10% glycerol (v/v), 0.05% Triton X-100 (v/v) and 1 mM 2-mercaptoethanol (BME). Lysozyme and DTT were added to a final concentration of 25 μg mL^{-1} and 5 mM respectively. Resuspended cells were incubated at room temperature for 20 minutes before sonication at 50 Hz and cell disruption at 20.1 kpsi. DNAse I was added to a final concentration of 5 μg mL^{-1} and allowed to incubate at room temperature for 20 minutes. Cell free extract was centrifuged at 20300 x g at 4º C for 15 minutes and the supernatant was filtered through a 0.2 μm filter. For assays involving cell free extract, the cell free extract was used without any further purification.

To the filtered cell free extract was then added 1, 10 phenanthroline to a final concentration of 1 mM and the cell free extract was incubated on ice for 15 minutes. The cell free extract solution was added to 2 mL of GST resin (GE Life Science) prepared according to the manufacturer's instructions. Cell free extract was passed through the column a total of three times. The column was washed with PBS buffer (25 mL) with 0.05% Triton X-100 (v/v), 10 % glycerol (v/v) and 1 mM BME before the protein was eluted into GST elution buffer (10 mL) plus 20 mM reduced glutathione, 10 % (v/v) glycerol (v/v), 1 mM BME and 0.05% Triton X-100 (v/v). Protein was then concentrated using a 50 kDa centrifugal filter unit before buffer exchange using a PD10 gel filtration column (GE Life Sciences) according to the manufacturer's instructions. Protein was eluted into PBS buffer (3.5 mL) plus 10% glycerol (v/v), 1 mM BME and 0.05% Triton X-100 (v/v) and concentrated further in a 10 kDa centrifugal filter. Protein concentration was determined via the method of Bradford (Bradford 1976) and protein was analysed via SDS-PAGE using an 8% gel. Protein was aliquoted, snap frozen in liquid nitrogen and stored at -80º C.

8.6.3 Western Blot Analysis of GST-ZmNCED1

Following SDS-PAGE, wet transfer from the polyacrylamide gel to Hybond-P PVDF nitrocellulose membrane (prepared according to the manufacturer's instructions (GE

Life Sciences)) was performed for one hour at 4º C at 100 V in wet transfer buffer. Following transfer, the membrane was blocked in PBST buffer plus 5% milk powder (w/v) for one hour at room temperature. The membrane was washed twice in PBST before incubation with PBST buffer with 0.02% anti-GST-HRP-conjugate antibody (v/v) for one hour at room temperature. The membrane was finally washed briefly with PBST. Chemiluminescence was performed using the Amersham ECL Prime Western blotting detection kit (GE Life Science) and visualised using a BioRad ChemiDoc system

8.6.4 Purification of 9'-cis-Neoxanthin

Fresh organic spinach (≈ 100 g, Tesco) was added to methanol (200 mL) with sodium bicarbonate (0.4 g) and butylated hydroxytoluene (0.04 g). Spinach was homogenised with a hand blender for five minutes and the resulting solution was filtered through a sieve (≈2 mm pore) to give a dark green solution, which was subsequently filtered under suction. The filtrate was partitioned between cold saturated sodium chloride (100 mL) and cold diethyl ether (200 mL). The ether layer was extracted and the organic layer was washed again with cold diethyl ether (200 mL). Combined ether extracts were washed with cold sodium chloride (100 mL) and concentrated under reduced pressure to give a dark green oil resin. Alumina (10 g) was activated by the addition of distilled water (1.0 mL) and petroleum ether (20 mL) and allowed to stir for 5 minutes at room temperature. Activated alumina was used to prepare a column and was subsequently washed with petroleum ether (50 mL). Resin was resuspended in diethyl ether (2 mL) and petroleum ether (10 mL) and applied to the column. The column was eluted with 1 : 1 petroleum ether : diethyl ether (100 mL), diethyl ether (200 mL) and 5% ethanol in diethyl ether (100 mL) which finally eluted the 9'-*cis*-neoxanthin. Triton X-100 was added to a final concentration of 0.1% (v/v) and solvent was removed under reduced pressure to give a yellow solid which was resuspended in ethanol (1 mL) and stored in the dark at -20º C. Data: λ_{MAX} 416, 440 and 467 nm; HRMS (MicroTOF) 601.4251, calc. 601.4244 for $C_{40}H_{56}O_4$. ε at 440 nm: 1.33 x 10^5 M^{-1} cm^{-1}.

8.6.5 HPLC Assays of LeNCED1 and ZmNCEDs

For assays involving purified apo-LeNCED, the apo-protein was allowed to defrost on ice before 1 mM iron (II) sulfate and 1 mM sodium ascorbate were added. The protein sample was then allowed to incubate on ice for a further 10 minutes before being assayed.

For assays involving cell free extract, the cell free extract was pre-incubated with the addition of 1.3 mM iron (II) sulfate and 1.3 mM sodium ascorbate. The cell free extract was allowed to incubate on ice for 5 minutes before addition to the assay.

To NCED assay buffer (130 or 150 μL) was added fresh 9'-*cis*-neoxanthin extract (25 μL, ≈80 μM), cell lysate containing overexpressed NCED enzyme (50 μL) or purified LeNCED1 (20 μL, ≈50 μg). Samples were incubated in the dark at room temperature for 15 minutes. Water (700 μL) was added and the products were extracted into ethyl acetate (3 × 700 μL). Ethyl acetate was filtered through cotton wool and removed under nitrogen. The residue was dissolved in methanol (300 μL) and 50 μL was injected into a Phenomenex HyperClone C_{18} reverse phase HPLC column (5μ BDS, 130Å, 250 x 4.6 mm) maintained at room temperature at a flow rate of 0.5 mL min^{-1}. The column was eluted with 95:5 acetonitrile : methanol, shifting to 85:10:5 acetonitrile : methanol : ethyl acetate over 20 minutes. Over the next 5 minutes, the gradient was shifted back to 95:5 acetonitrile : methanol and held for a further 20 minutes. C_{25} product and unreacted 9'-*cis*-neoxanthin were observed at 440 nm. Retention times: 9'-*cis*-neoxanthin, 10.2 minutes; C_{25} product, 6.5 minutes.

Inhibitors were dissolved in either ethanol or DMSO and a total of 2 μL of inhibitor stock was added to the assay to give a final concentration of 100 μM. For time dependent inhibition assays, inhibitor was added to assay mix containing LeNCED enzyme for 10 minutes before the addition of substrate. For other assays, enzyme was added to a mix of substrate and inhibitor and water was added to control assays in place of inhibitor. Inhibition was calculated from the product formation after 30 minutes, compared with a control assay with no inhibitor.

8.7 *Arabidopsis thaliana* ABA2

8.7.1 Cloning of AtABA2

Arabidopsis thaliana ABA2 DNA (sequence optimised) was purchased from Genscript cloned into the pUC-57-AtABA2 plasmid. Primers for the AtABA2 gene were designed, incorporating a CACC overhang on the 5' terminus for ligation into the D-TOPO pET-200 vector (Invitrogen). The AtABA2 gene was amplified by PCR from pUC57-AtABA2 using the following primers: Forward 5' CAC CAT GTC AAC GAA TAC CGA ATC CTCT 3'; Reverse 5' TCA GCG AAA CAC TTT GAA CGA ATG 3'. PCR: 1 cycle of 94° C for 120 seconds; 35° C cycles of 94° C for 30 seconds, 58° C for 30 seconds, 72° C for 60 seconds; 1 cycle of 72° C for 120 seconds. PCR reaction consisted of water (39.1 μL), 10X *pfx* buffer (Invitrogen) (5.0 μL), 50 mM magnesium sulfate (1.0 μL), 10 mM dNTPs (1.5 μL), 10 μM forward primer (1.5 μL), 10 μM reverse primer (1.5 μL), platinum *pfx* tag polymerase (1.0 μL) and pUC57-AtABA2 (1.0 μL, 0.6 μg). PCR product was separated by agarose gel electrophoresis and visualised under UV-light at 260 nm. AtABA2 was excised and purified using the Fermentas DNA extraction kit according to the manufacturer's instructions.

AtABA2 with CACC overhang was ligated into the D-TOPO pET-200 vector according to the manufacturer's instructions. Plasmid was inserted into TOP10 *E. coli* (Invitrogen) according to the procedure in Section 8.4.2. Colonies were screened by sequencing and PCR using primers for AtABA2 gene amplification. Plasmid was extracted from an overnight culture (10 mL) using the Fermentas plasmid extraction kit according to the manufacturer's instructions and stored at -20° C.

pET-200 AtABA2 plasmid was used to transform BL21 *E. coli* according to the procedure detailed in Section 8.4.2.

8.7.2 Expression and Purification of AtABA2

Cells were prepared as described in section 8.4.3 using pET-200-AtABA2 *E. coli* BL21 and were resuspended in N-His_6 wash buffer (5 mL) with 25 μg μL^{-1} lysozyme and

100 mM PMSF. Cells were lysed by cell disruption at 20.1 kpsi. Cellular debris was pelleted by centrifugation at 20300 x g at 4º C for 15 minutes before the lysate was syringe filtered through a 0.2 µm filter. Nickel resin (1 mL, GE Life Science) was prepared according to the manufacturer's instructions. Cell free extract containing overproduced N-His_6-AtABA2 was loaded onto the nickel resin and allowed to incubate at 4º C for one hour. Resin was added to a column and washed with N-His_6 wash buffer (20 mL). AtABA2 was eluted into N-His_6 elution buffer (5 mL) before buffer exchange into N-His_6 wash buffer using a PD10 column (GE Life Science). AtABA2 was then further purified by FPLC. AtABA2 was loaded onto a GE Healthcare His tag column (1 mL) at a flow rate of 0.5 mL min^{-1}. Column was washed with N-His_6 wash buffer (20 mL) at a flow rate of 0.5 mL min^{-1}. Finally, protein was eluted into N-His_6 elution buffer (10 mL) at a flow rate of 1 mL min^{-1}. Fractions were analysed by SDS-PAGE on a 15% gel and protein was concentrated in a 10 kDa centrifugal filter unit before buffer exchanged into 20 mM MOPS buffer. AtABA2 was aliquoted into 200 µL aliquots and stored at -80° C until required. Protein was quantified via the method described by Bradford (Bradford 1976).

8.7.3 Assays of AtABA2

AtABA2 assays were monitored continuously via UV at 340 nm. Assays were conducted in a total volume of 1000 µL in potassium phosphate buffer with 100 µM nicotinamide adenine dinucleotide (NAD^+) and 30 µg mL^{-1} AtABA2. In the absence of authentic xanthoxin substrate, 3,3,5-trimethylcyclohexanol was used at a final concentration of 100 µM.

Control assays for ABA2 using fluorescence were conducted in a total volume of 200 µL potassium phosphate buffer using a microtiter plate reader using the same concentration of reagents as for UV assays. Samples were excited at 340 nm and the emission at 440 nm was observed.

8.7.4 Coupled Assays of AtABA2 and NCEDs

Coupled assays were conducted in a total volume of either 1000 µL or 200 µL in potassium phosphate buffer with 100 mM NAD^+, 30 µg mL^{-1} ABA2, 20-50 µL of cell free extract containing overproduced NCED or 20 µL of purified NCED (≈40 µg) and 40 µM 9'-*cis*-neoxanthin. Reactions were monitored continuously at 340 nm.

Coupled assays using luminescence were conducted using the Promega NAD(P)H Glo kit in a 100 µL total volume in a microtiter plate and were monitored using a microtiter plate reader. 50 µL of a normal coupled assay mix following the reaction was added to 50 µL of luminescent reagent and the assay was conducted according to the manufacturer's instructions.

8.8 *Oryza Sativa* D27

8.8.1 Cloning of OsD27.1 and OsD27.2 into pET-151

Oryza sativa OsD27.1 (full length sequence, sequence optimised) was purchased from Genscript cloned into the pUC-57-KAN-OsD27 plasmid. Primers for the full length OsD27 gene (OsD27.1) and truncated peptide lacking the first 120 base pairs corresponding to a transit peptide (OsD27.2) were designed, incorporating a CACC overhang on the 5' terminus for ligation into the D-TOPO pET-151 vector (Invitrogen) (Table 8.4). The OsD27

Construct		Primer Sequence (5'-3')
pET-151-OsD27.1	F	CAC CAT GGA AAC GAC GAC CCT GGT TC
	R	TCA AAT GCT ACA GTT CAC GCC ATG
pET-151-OsD27.2	F	CAC CGC AGT GAT GGC ACG CCC G
	R	TCA AAT GCT ACA GTT CAC GCC
pMAL-c5e-OsD27.1	F	**GAT ATC** ATG GAA ACG ACG ACC C
	R	**GGA TCC** TCA AAT GCT ACA GTT C
pGEX-4T-1-OsD27.1	F	ATC AGT **GGA TCC** ATG GAA ACG ACG ACC CTG GTT C
	R	TAG ACT **GAA TTC** TCA AAT GCT ACA GTT CAC GCC ATG

Table 8.4 – Primer sequences used for cloning of OsD27 into pET-151, pMAL-c5e and pGEX-4T-1 vectors.

gene was amplified by PCR from pUC-57-KAN-OsD27 using the primers shown in Table 8.4. PCR: 1 cycle of 94° C for 120 seconds; 35° C cycles of 94° C for 30 seconds, 58° C for 30 seconds, 72° C for 60 seconds; 1 cycle of 72° C for 120 seconds. PCR reaction consisted of water (39.1 μL), 10X *pfx* buffer (Invitrogen) (5.0 μL), 50 mM magnesium sulfate (1.0 μL), 10 mM dNTPs (1.5 μL), 10 μM forward primer (1.5 μL), 10 μM reverse primer (1.5 μL), platinum *pfx* tag polymerase (1.0 μL) and pUC-57-KAN-OsD27 (1.0 μL, 0.6 μg). PCR products were separated by agarose gel electrophoresis and visualised under UV-light at 260 nm. OsD27 was excised and purified using the Fermentas DNA extraction kit according to the manufacturer's instructions.

OsD27 with GACC overhang was ligated into the D-TOPO pET-151 vector (Invitrogen) according to the manufacturer's instructions. Plasmid was inserted into chemically competent TOP10 *E. coli* (Invitrogen) according to the procedure in Section 8.4.2. Colonies were screened by sequencing and PCR using primers for pET-151-OsD27 gene amplification (Table 8.4). Plasmid was extracted from an overnight culture (10 mL) using the Fermentas plasmid extraction kit according to the manufacturer's instructions and the plasmid was stored at -20° C.

pET-151-OsD27 plasmid was used to transform pRoestta pBB541 BL21 *E. coli* according to the procedure detailed in Section 8.4.2.

8.8.2 Cloning of OsD27.1 into pMAL-c5e and pGEX-4T-1 Vector

Oryza sativa OsD27.1 (full length sequence, sequence optimised) was purchased from Genscript cloned into the pUC-57-KAN-OsD27 plasmid. The OsD27 gene was amplified by PCR from pUC-57-KAN-OsD27 using the primers shown in Table 8.4. PCR: 1 cycle of 94° C for 300 seconds; 35° C cycles of 94° C for 30 seconds, 55° C for 30 seconds, 72° C for 60 seconds; 1 cycle of 72° C for 120 seconds. PCR reaction consisted of water (39.1 μL), 10X *pfx* buffer (Invitrogen) (5.0 μL), 50 mM magnesium sulfate (1.0 μL), 10 mM dNTPs (1.5 μL), 10 μM forward primer (1.5 μL), 10 μM reverse primer (1.5 μL), platinum *pfx* tag polymerase (1.0 μL) and pUC-57-KAN-OsD27 (1.0 μL, 0.6 μg). PCR product was

separated by agarose gel electrophoresis and visualised under UV-light at 260 nm. OsD27 was excised and purified using the Fermentas DNA extraction kit according to the manufacturer's instructions.

OsD27.1 PCR product (0.2 µg), pGEX-4T-1 (1 µg) (GE Life Sciences) and pMAL-c5e vector (1 µg) (New England Biolabs) were digested with the *Bam*HI and *Eco*RV (for pMAL) or *Bam*HI and *Eco*RI (for pGEX) (Thermo Scientific FastDigest) restriction enzymes (1 µL each) in distilled water (to a total volume of 40 µL) and FastDigest green buffer (4 µL), incubated at 37° C for one hour. Digestion products were purified via agarose gel electrophoresis and extracted using the Fermentas DNA extraction kit according to the manufacturer's instructions.

Digested PCR products and vector were ligated with T4 DNA ligase (Invitrogen). To vector (pGEX-4T-1 or pMAL-c5e) (20 ng) was added PCR product (100 ng), 10 X ligation buffer (Invitrogen, 2 µL), T4 ligase (1 µL) and water (to a total volume of 20 µL). The mixture was incubated at 4° C overnight. Ligation mixture (3 µL) was transformed into chemically competent *E. coli* DH5α according to the procedure detailed in section 8.4.2. Colonies were screened by sequencing and PCR. Plasmid was extracted from an overnight culture (10 mL) using the Fermentas plasmid extraction kit according to the manufacturer's instructions and the plasmid was stored at -20° C.

pMAL-c5e-OsD27.1 and pGEX-4T-1 plasmids were used to transform pRosetta pBB541 BL21 or BL21 *E. coli* according to the procedure detailed in Section 8.4.2.

8.8.3 Preparation of Cell Free Extract and Purification of OsD27 from pET vector

Cells were prepared as described in section 8.4.3 using pET-151-OsD27(.1 or .2) *E. coli* and were resuspended in N-His_6 wash buffer (5 mL) with 0.05% Triton X-100 (v/v), 5 mM DTT and 100 mM PMSF. Cells were lysed by cell disruption at 20.1 kpsi. DNAse I was added to a final concentration of 5 µg mL^{-1} and allowed to incubate at room temperature for 20 minutes. Cellular debris was pelleted by centrifugation at 20300 x g at 4º c for 15

minutes before the lysate was syringe filtered through a 0.2 µm filter. Cell free extract containing overproduced N-His_6-OsD27 was loaded onto a HisTrap Gravitrap column (1 mL, GE Life Science) for a total of three passages. The column was washed with N-His_6 wash buffer (20 mL). OsD27 was eluted into N-His_6 elution buffer (5 mL). Protein was then concentrated using a 10 kDa centrifugal filter unit before buffer exchange using a PD10 gel filtration column (GE Life Sciences) used according to the manufacturer's instructions. Protein was eluted into 100 mM bis-tris buffer pH 6.7 (3.5 mL) plus 10% glycerol (v/v) and concentrated further in a 10 kDa centrifugal filter. Protein concentration was determined via the method of Bradford (Bradford 1976) and protein was analysed via SDS-PAGE using a 10% gel. Protein was aliquoted, snap frozen in liquid nitrogen and stored at -80º C.

8.8.4 Preparation of Cell Free Extract and Purification of OsD27 from pMAL vector

Cells were prepared as described in section 8.4.3 using pMAL-c5e-OSD27.1 *E coli* and were resuspended MBP wash buffer (5 mL) with 0.05% Triton X-100 (v/v), 5 mM DTT and 100 mM PMSF. Cells were lysed by cell disruption at 20.1 kpsi. DNAse I was added to a final concentration of 5 µg mL^{-1} and allowed to incubate at room temperature for 20 minutes. Cellular debris was pelleted by centrifugation at 20300 x g at 4º C for 15 minutes before the lysate was syringe filtered through a 0.2 µm filter. Cell free extract containing overproduced MBP-OsD27 was loaded onto amylose resin (2 mL, New England Biolabs, prepared according to the manufacturer's instructions) and incubated at 4° C for two hours. Resin was washed with MBP wash buffer (20 mL) and OsD27 was eluted into MBP elution buffer (5 mL) before buffer exchange into N-His_6 wash buffer. Protein was then concentrated using a 10 kDa centrifugal filter unit before buffer exchange using a PD10 gel filtration column (GE Life Sciences) used according to the manufacturer's instructions. Protein was eluted into 100 mM bis-tris buffer pH 6.7 (3.5 mL) plus 10% glycerol (v/v) and concentrated further in a 10 kDa centrifugal filter. Protein concentration was determined via the method of Bradford (Bradford 1976) and analysed via SDS-PAGE using a 10% gel.

8.8.5 Preparation of Cell Free Extract and Purification of OsD27 from pGEX vector

Cells were prepared as described in Section 8.4.3 with pGEX-4T-1-OsD27.1 and were resuspended into PBS buffer (10 mL) with 0.05% Triton X-100 (v/v), 1 mM PMSF and 5 mM DTT. Resuspended cells were incubated at room temperature for 20 minutes before sonication at 50 Hz and cell disruption at 20.1 kpsi. DNAse I was added to a final concentration of 5 µg mL^{-1} and allowed to incubate at room temperature for 20 minutes. Cell free extract was centrifuged at 20300 x g at 4º C for 20 minutes and the supernatant was filtered through a 0.2 µm filter.

The cell free extract solution was added to a 2 mL GST-Gravitrap column (GE Life Science) prepared according to the manufacturer's instructions. Cell free extract was passed through the column a total of three times. The column was washed with PBS buffer (25 mL) and eluted with GST elution buffer (10 mL) plus 20 mM reduced glutathione. Protein was then concentrated using a 10 kDa centrifugal filter unit before buffer exchange using a PD10 gel filtration column (GE Life Sciences) used according to the manufacturer's instructions. Protein was eluted into 100 mM bis-tris buffer pH 6.7 (3.5 mL) plus 10% glycerol (v/v) and concentrated further in a 10 kDa centrifugal filter. Protein concentration was determined via the method of Bradford (Bradford 1976) and protein was analysed via SDS-PAGE using an 8% gel. Protein was aliquoted, snap frozen in liquid nitrogen and stored at -80º C.

8.8.6 Western Blot Analysis of MBP-OsD27.1 and GST-OsD27.1

Following SDS-PAGE, wet transfer from the polyacrylamide gel to Hybond-P PVDF nitrocellulose membrane (prepared according to the manufacturer's instructions (GE Life Sciences)) was performed for one hour at 4º C at 100 V in wet transfer buffer. Following transfer, the membrane was blocked in PBST buffer plus 5% milk powder (w/v) for one hour at room temperature. The membrane was washed twice in PBST before incubation with PBST buffer with 0.02% anti-GST-HRP-conjugate antibody (for GST-OsD27.1, GE Life Sciences) or anti-MBP-HRP-conjugate antibody (for MBP-OsD27, New England Biolabs) (v/v) for one hour at room temperature. The membrane was finally washed

briefly with PBST. Chemiluminescence was performed using the Amersham ECL Prime Western blotting detection kit (GE Life Science) and visualised via autoradiography onto Amersham Hyperfilm ECL (GE Life Sciences) and developed with Carestream GBX developer and fixer solutions (Kodak) according to the manufacturer's instructions.

8.8.7 Isomerisation of β-carotene

To a solution of *β*-carotene (20 mg) dissolved in hexane (20 mL) was added 1 mL of iodine (10 mg) in hexane (10 mL) solution. The resulting iodine-*β*-carotene-hexane solution was exposed to UV light (210 nm) for 30 minutes at room temperature.

9-*cis*-*β*-carotene (100 µL injection) was purified via HPLC using a Thermo Scientific Acclaim C_{30} reversed phase column (5 µm, 4.6 x 250 mm) maintained at 4° C at a flow rate of 1.0 mL min^{-1}. Column was eluted with 75 : 25 methanol : acetonitrile, shifting to 25 : 35 : 50 acetonitrile : methanol : tertbutylmethylether (TBME) over 20 minutes. Over the next 5 minutes, the gradient was held at 25 : 35 : 50 acetonitrile : methanol : TBME before returning to 75 : 25 methanol : acetonitrile for a final five minutes. 9-*cis*-*β*-carotene was observed at 475 nm. Data: λ_{MAX} 448, 474 nm. ε at 475 nm: 6.55 x 10^4 M^{-1} cm^{-1}. Retention time of 9-*cis*-*β*-carotene: 24.5 minutes.

8.8.8 In vitro Assays of OsD27

Assays were conducted in a total volume of 200 µL. D27 enzyme (10 µL, 40 µg) was preincubated with 1 µM iron (II) sulphate and 1 µm 2-mercaptoethanol for 10 minutes on ice prior to the addition of D27 assay buffer (to a total volume of 200 µL) and 9-*cis* or all-*trans*-*β*-carotene (20 µM in ethanol). Assays were incubated at 25° C with shaking at 180 rpm in the dark for 20 minutes. Assays were stopped with the addition of 4:1:4 acetone : petroleum ether : diethyl ether (400 µL). The resulting solution was centrifuged at 4220 x g for five minutes. The epiphase was collected and removed in a desiccator under reduced pressure. The residue was resuspended in methanol (200 µL) and analysed by HPLC (using the method in Section 8.8.7).

Inhibitors were dissolved in either ethanol or DMSO and a total of 2 µL of inhibitor stock was added to the assay to give a final concentration of 100 µM. Inhibitor was added to assay mix containing enzyme for 10 minutes before the addition of substrate. Water was added to control assays in place of inhibitor. Inhibition was calculated from peak areas determined by integration of peaks after 5 minutes, compared with a control assay with no inhibitor.

8.8.9 Mechanistic Assays of OsD27

Assays were conducted using the same conditions as described in Section 8.8.8. Hydrogen peroxide, 2-mercaptoethanol, sodium dithionite, sodium dithionite / pyridine, lead (II) acetate and silver (I) acetate were all used at a final concentration of 100 µM and assays performed as described in Section 8.8.8.

pH rate profile experiments were performed using the same conditions as Section 8.8.8 with the exception that D27 buffer was replaced. For pH 4.12, 4.65, 5.13, 5.61 and 5.94 100 mM sodium acetate buffer was used, for pH 6.71 100 mM PBS buffer was used and for pH 7.70, 8.19, 8.62, 9.0, 9.62 and 10.3 100 mM tris-HCl was used.

For oxygen dependence assays, assays were set up as described in Section 8.8.8 in a two valved vessel. The system was thoroughly degassed and flushed three times with nitrogen. Reactions were performed under an atmosphere of nitrogen.

UV assays were performed in a total volume of 50 µL in an Eppendorf UVette containing approximately 50 µg of protein in 100 mM bis-tris buffer plus 10% glycerol (v/v) and 0.1% Triton X-100 (v/v). Hydrogen peroxide, 2-mercaptoethanol, sodium dithionite, sodium dithionite / pyridine, lead (II) acetate and silver (I) acetate were all used at a final concentration of 100 µM.

8.8.10 In vivo Assays of OsD27

pGEX-4T-1-OsD27.1 and pET-151-OsD27.1 were used to transform competent pAC-BETA BL21 *E. coli* (Section 8.4.1 and 8.4.2). LB media (10 mL) containing

appropriate the antibiotic (see Table 8.2) was inoculated with a single colony of pGEX-4T-1-OsD27.1 / pET-151-OsD27.1 pAC-BETA *E. coli* BL21 (containing desired plasmid(s)) and grown overnight at 37º C, 180 rpm. LB media (200 mL) containing appropriate antibiotics was inoculated with the overnight culture and grown at 37º C, 180 rpm until OD at 600 nm = 0.6 AU. IPTG was added to a final concentration of 1 mM. Cultures were incubated overnight at 20º C, 180 rpm, harvested at 4220 x g at 4º C for 15 minutes and then resuspended in water (5 mL). Lysozyme was added to give a final concentration of 25 µg mL^{-1}. Following incubation at room temperature for 10 minutes, cells were lysed by sonication at 50 Hz. *n*-Butanol (1 mL) was added and the organic phase was collected following centrifugation at 4220 x g at 4º C for 10 minutes. The butanol was added to a 100 mg 1 mL Agilnet Bond Elut C_{18} solid phase extraction column and centrifuged at 2000 x g at 4º C for 2 minutes. The column was washed with 10 x 1 mL water before the *β*-carotene was eluted into TBME (1 mL). TBME was removed under reduced pressure and the residue resuspended into 200 µL methanol. Samples were analysed via HPLC using a Thermo Scientific Acclaim C_{30} reversed phase column using the same method described in Section 8.8.7.

For inhibition assays, D4 was added upon induction with IPTG and inhibitor was added to a final concentration of 100 µM.

8.8.11 Coupled OsD27 / AtCCD7 Assays

Assays were conducted in a total volume of 200 µL. OsD27.1 enzyme was pre-incubated with 1 µM iron (II) sulphate and 1 µm 2-mercaptoethanol for 10 minutes on ice. OsD27.1 (10 µL, 40 µg) was then added to all-*trans*-*β*-carotene (20 µM in ethanol) and AtCCD7.2 (10 µL, 30 µg) Assays were incubated at 25° C with shaking at 180 rpm in the dark for 20 minutes. Assays were stopped with the addition of 4:1:4 acetone : petroleum ether : diethyl ether (400 µL). The resulting solution was centrifuged at 4220 x g for five minutes. The epiphase was collected and removed in a desiccator under reduced pressure. The residue was resuspended in methanol (200 µL) and analysed by HPLC (Section 8.8.7).

Inhibitors (in DMSO or ethanol) were added to a final concentration of 100 µM and pre-incubated with OsD27.1 and AtCCD7.2 before the addition of all-*trans*-*β*-carotene.

8.9 *Arabidopsis thaliana* CCD7

8.9.1 Preparation of E. coli BL21 Containing pBB541 and pGEX-4T-1-AtCCD7 Plasmids

Chemically competent pRosetta *E. coli* BL21 was purchased from Millipore. The pBB541 plasmid (Addgene) was transformed into chemically competent pRosetta BL21 via the method in Section 8.4.2. Chemically competent pBB541 pRosetta *E. coli* BL21 was prepared from the transformants using the method in Section 8.4.2 and the pGEX-4T-1-AtCCD7.1 / pGEX-4T-1-AtCCD7.2 plasmids were transformed into the chemically competent pBB541 pRosetta *E. coli* BL21 using the method in Section 8.4.2.

8.9.2 Preparation of Cell Free Extract and Purification of AtCCD7.1/.2

Cells were prepared as described in Section 8.4.3 with the appropriate *E. coli* strain and were resuspended into PBS buffer (10 mL) with 0.05% Triton X-100 (v/v), 1 mM PMSF and 5 mM DTT. Resuspended cells were incubated at room temperature for 20 minutes before sonication at 50 Hz and cell disruption at 20.1 kpsi. DNAse I was added to a final concentration of 5 µg mL^{-1} and allowed to incubate at room temperature for 20 minutes. Cell free extract was centrifuged at 20300 x g at 4º C for 20 minutes and the supernatant was filtered through a 0.2 µm filter.

The cell free extract solution was added to a 2 mL GST-Gravitrap column (GE Life Science) prepared according to the manufacturer's instructions. Cell free extract was passed through the column a total of three times. The column was washed with PBS buffer (25 mL) and eluted into GST elution buffer (10 mL) plus 20 mM reduced glutathione. Protein was then concentrated using a 10 kDa centrifugal filter unit before buffer exchange using a PD10 gel filtration column (GE Life Sciences) used according to the manufacturer's instructions. Protein was eluted into 100 mM bis-tris buffer pH 6.7 (3.5 mL) plus 10% glycerol (v/v) and concentrated further in a 10 kDa centrifugal filter. Protein concentration was determined via

the method of Bradford (Bradford 1976) and protein was analysed via SDS-PAGE using an 8% gel. Protein was aliquoted, snap frozen in liquid nitrogen and stored at -80º C.

8.9.3 Western Blot Analysis of AtCCD7.2

Western blotting was performed as described in Section 8.8.6.

8.9.4 Assays of AtCCD7.2

CCD7 enzyme (preincubated with 1 µM iron (II) sulphate and 1 µm sodium ascorbate for 10 minutes on ice) (10 µL, 30 µg) was added to 9-*cis* or all-*trans*-*β*-carotene (20 µM) with CCD7 assay buffer (to a total volume of 200 µL). Assay were incubated at 25° C with shaking at 180 rpm in the dark for 20 minutes. Assays were stopped with the addition of 4 M pyridinium acetate or 4 M sodium chloride (200 µL) and ethyl acetate (200 µL). The resulting solution was centrifuged at 4220 x g for five minutes. The epiphase was collected and added to a 100 mg 1 mL C_{18} solid phase extraction column (Agilent Bond Elut) and centrifuged at 2000 x g. Columns were washed with 10 x 1 mL of distilled water and the products eluted with TBME (600 µL). TBME was removed in a desiccator under reduced pressure. The residue was resuspended in methanol (200 µL) and analysed by HPLC (100 µL injection) using the method in Section 8.8.7. 9-*cis*-*β*-apo-10'-carotenal elution at 10.5 minutes. Data: λ_{MAX} 444 nm; HRMS (MicroTOF) 377.2839, calc. 377. 2840 for $C_{27}H_{36}O+H^+$.

Inhibitors were dissolved in either ethanol or DMSO and a total of 2 µL of inhibitor stock was added to the assay to give a final concentration of 100 µM. Inhibitor was added to assay mix containing enzyme for 10 minutes before the addition of substrate. Water was added to control assays in place of inhibitor. Inhibition was calculated from peak areas determined by integration of peaks after 20 minutes, compared with a control assay with no inhibitor.

8.10 *Arabidopsis thaliana* CCD8

8.10.1 Cloning of AtCCD8.2 into pET-151

Full length AtCCD8 (CCD8.1) was obtained from Dr Andrew Thompson (Cranfield University). pGEX-4T-1-AtCCD8.1 was extracted from a 10 mL overnight culture of pGEX-4T-1-AtCCD8.1 DH5α *E. coli* using a Thermo Scientific plasmid extraction kit. Primers for the AtCCD8.2 gene were designed, incorporating a CACC overhang on the 5' terminus for ligation into the D-TOPO pET-151 vector. The AtCCD8.2 gene was amplified by PCR from pGEX-4T-1-AtCCD8.1 using the following primers: Forward 5' CAC CGC TGT AAT TAA TTC AGC GGC ACC 3'; Reverse 5' TCA ATC TTT GGG GAT CCA GCA AC 3'. PCR: 1 cycle of 94° C for 120 seconds; 35° C cycles of 94° C for 30 seconds, 58° C for 30 seconds, 72° C for 60 seconds; 1 cycle of 72° C for 120 seconds. PCR reaction consisted of water (39.1 μL), 10X *pfx* buffer (Invitrogen) (5.0 μL), 50 mM magnesium sulfate (1.0 μL), 10 mM dNTPs (1.5 μL), 10 μM forward primer (1.5 μL), 10 μM reverse primer (1.5 μL), platinum *pfx* tag polymerase (1.0 μL) and pUC57-AtABA2 (1.0 μL, 0.6 μg). PCR product was separated by agarose gel electrophoresis and visualised under UV-light at 260 nm. AtCCD8.2 was excised and purified using the Fermentas DNA extraction kit according to the manufacturer's instructions.

AtCCD8.2 with CACC overhang was ligated into the D-TOPO pET-151 vector (Invitrogen) according to the manufacturer's instructions. Plasmid was inserted into TOP10 *E. coli* (Invitrogen) according to the procedure in Section 8.4.2. Colonies were screened by sequencing and PCR using primers for AtCCD8.2 gene amplification. Plasmid was extracted from an overnight culture (10 mL) using the Fermentas plasmid extraction kit according to the manufacturer's instructions and the plasmid was stored at -20° C.

8.10.2 Preparation of pET-151-AtCCD8.2 pRosetta E. coli BL21

Chemically competent pRosetta *E. coli* BL21 was purchased from Millipore. pET-151-AtCCD8.2 was extracted from a 10 mL overnight culture of pET-151-AtCCD8.2 DH5α *E. coli* using the Thermo Scientific plasmid extraction kit and the pET-151-AtCCD8.2

plasmid was used to transform chemically competent pRosetta *E. coli* BL21 using the method in Section 8.4.2.

8.10.3 Preparation of pGEX-4T-1-AtCCD8.1 E. coli BL21

Chemically competent *E. coli* BL21 was prepared according to the method in Section 8.4.1. The pGEX-AtCCD8.1 was extracted from a 10 mL overnight culture of pGEX-4T-1-AtCCD8.1 DH5α *E. coli.* The pGEX-4T-1-AtCCD8.1 plasmid was used to transform chemically competent BL21 via the method in Section 8.4.2.

8.10.4 Preparation of pGEX-4T-1-AtCCD8.2 pRosetta pBB541 E. coli BL21

Chemically competent pRosetta *E. coli* BL21 was purchased from Millipore. The pBB541 plasmid was used to transform chemically competent BL21 via the method in Section 8.4.2. Chemically competent pRosetta pBB541 *E. coli* BL21 was prepared from the transformants using the method in Section 8.4.2. pGEX-AtCCD8.2 was extracted from a 10 mL overnight culture of pGEX-4T-1-AtCCD8.2 DH5α *E. coli* using the Thermo Scientific plasmid extraction kit and the pGEX-4T-1-AtCCD8.2 plasmid was used to transform the chemically competent *E. coli* BL21 pRosetta pBB541 using the method in Section 8.4.2.

8.10.5 Preparation of Cell Free Extract and Purification of GST-AtCCD8.1/.2 from pGEX vector

Cells were prepared as described in Section 8.4.3 with the appropriate *E. coli* strain and were resuspended into PBS buffer (10 mL) with 0.05% Triton X-100 (v/v), 1 mM PMSF and 5 mM DTT. Resuspended cells were incubated at room temperature for 20 minutes before sonication at 50 Hz and cell disruption at 20.1 kpsi. DNAse I was added to a final concentration of 5 μg mL^{-1} and allowed to incubate at room temperature for 20 minutes. Cell free extract was centrifuged at 20300 x g at 4º C for 20 minutes and the supernatant was filtered through a 0.2 μm filter.

The cell free extract solution was added to a 2 mL GST-Gravitrap column (GE Life Science) prepared according to the manufacturer's instructions. Cell free extract was passed through the column for a total of three passages. The column was washed with PBS buffer

(25 mL) and eluted into GST elution buffer (10 mL) plus 20 mM reduced glutathione. Protein was then concentrated using a 10 kDa centrifugal filter unit before buffer exchange using a PD10 gel filtration column (GE Life Sciences) used according to the manufacturer's instructions. Protein was eluted into 100 mM bis-tris buffer pH 6.7 (3.5 mL) plus 10% glycerol (v/v) and concentrated further in a 10 kDa centrifugal filter. Protein concentration was determined via the method of Bradford (Bradford 1976) and protein was analysed via SDS-PAGE using an 8% gel. Protein was aliquoted, snap frozen in liquid nitrogen and stored at -80º C.

8.10.6 Preparation of Cell Free Extract and Purification of AtCCD8 from pET Vector

Cells were prepared as described in section 8.4.3 using pET-151-AtCCD8.2 pRosetta *E. coli* and was resuspended in N-His_6 wash buffer (5 mL) with 0.05% Triton X-100 (v/v), 5 mM DTT and 100 mM PMSF. Cells were lysed by cell disruption at 20.1 kpsi. DNAse I was added to a final concentration of 5 µg mL^{-1} and allowed to incubate at room temperature for 20 minutes. Cellular debris was pelleted by centrifugation at 20300 x g at 4 ºc for 15 minutes before the lysate was syringe filtered through a 0.2 µm filter.

Cell free extract containing overproduced N-His_6-OsD27 was loaded onto a HisTrap Gravitrap column (1 mL, GE Life Science) and passed through the column a total of three times. Column and washed with N-His_6 wash buffer (20 mL). OsD27 was eluted into N-His_6 elution buffer (5 mL). Protein was then concentrated using a 10 kDa centrifugal filter unit before buffer exchange using a PD10 gel filtration column (GE Life Sciences) used according to the manufacturer's instructions. Protein was eluted into 100 mM bis-tris buffer pH 6.7 (3.5 mL) plus 10% glycerol (v/v) and concentrated further in a 10 kDa centrifugal filter. Protein concentration was determined via the method of Bradford (Bradford 1976) and protein was analysed via SDS-PAGE using a 10% gel. Protein was aliquoted, snap frozen in liquid nitrogen and stored at -80º C.

8.10.7 Assays of AtCCD8

The substrate for CCD8 assays, 9-*cis*-*β*-apo-10'-carotenal, were obtained from incubations of AtCCD7.2 with 9-*cis*-*β*-carotene via the same procedure as detailed in Section 8.8.4. 9-*cis*-*β*-apo-10'-carotenal was extracted from assays with 4 x 400 µL 4:1:4 acetone : petroleum ether : diethyl ether. The extracts were pooled and reduced under pressure. The residue was resuspended in ethanol and used either without further purification or was purified via the method detailed in Section 8.8.7. If the internal standard *β*-carotene was required, *β*-carotene (in ethanol) was added to a final concentration of 5 µM.

For HPLC assays, samples were performed in a total volume of 200 µL with CCD8 assay buffer. AtCCD8 (10 µL, 30 µg) was preincubated with 1 µM sodium ascorbate and 1 µM iron (II) sulphate for 10 minutes on ice before addition of 9-*cis*-*β*-apo-10'-carotenal (10 µL, 20 µM). Assays were incubated for 10 minutes in the dark at 25° C and were stopped with the addition of 4:1:4 acetone : petroleum ether : diethyl ether (400 µL). The resulting solution was centrifuged at 4220 x g for five minutes. The epiphase was collected and the solvent was removed in a desiccator under reduced pressure. The residue was resuspended in methanol (200 µL) and analysed by HPLC using the method in Section 8.8.7.

HPLC separations using the method of Alder *et al*. 2012 were performed using a Thermo Scientific Acclaim C_{30} reversed phase column (5 µm, 4.6 x 250 mm) maintained at room temperature at a flow rate of 1.0 mL min^{-1}. The column was eluted for 45 minutes shifting from 100% solvent system A, 0% solvent system B to 57% A, 43% B. Over the next minute, the gradient was shifted to 100% solvent system B and the gradient was maintained at 100% B for the next 34 minutes. Over the final 10 minutes the gradient was moved back to 100% A. Solvent system A: 30:10:1 methanol : water : TBME. Solvent system B: 1:1 methanol : TBME.

HPLC separations using the methanol/water gradient were performed a Phenomenex HyperClone C_{18} reverse phase HPLC column (5µ BDS, 130Å, 250 x 4.6 mm) maintained at room temperature at a flow rate of 0.5 mL min^{-1}. Column was eluted with 90% water, 10%

methanol shifting to 100% methanol over 30 minutes and maintained at 100% methanol for a further 20 minutes before the gradient was returned to 90% water, 10% methanol.

Inhibitors were dissolved in either ethanol or DMSO and incubated with enzyme at a final concentration of 100µM for one hour prior to addition of substrate. Inhibitor was also added to buffer such that upon addition of enzyme (containing inhibitor) to the assay mix (buffer and enzyme), the inhibitor was at a final concentration of 100 µM. Water was added to control assays in place of inhibitor. Inhibition was calculated from loss of 9-*cis*-*β*-apo-10'-carotenal relative to a *β*-carotene internal standard (5 µM).

For continuous UV assays reactions were performed in a 96 well microtitFre plate. Assays were performed in a total volume of 200 µL with CCD8 assay buffer. AtCCD8 was pre-incubated with 1 µM sodium ascorbate and 1 µM iron (II) sulphate for 10 minutes on ice 9-*cis*-*β*-apo-10'-carotenal (10 µL, 20 µM) was added to CCD8 buffer for 10 minutes before AtCCD8 (10 µL, 30 µg) was added. Reactions were monitored at 430 nm for 20 minutes.

Sodium borohydride reduction was performed by adding 5 mg of sodium borohydride to 100 µL 100 mM sodium hydroxide solution. 10 µL of this solution was added to 200 µL of organic extract from the CCD8 reaction. Analysis was performed by ESI-MS.

8.10.8 Biochemical Characterisation of AtCCD8

Stopped flow experiments were performed at the University of Manchester Institute of Biotechnology using an Applied Photophysics SX.18MV stopped flow spectrometer with a 05169 pbp monochromator. 9-*cis*-*β*-apo-10'-carotenal substrate was prepared as described in Section 8.9.7, with the exception that prior to the removal of 4:1:4 acetone : petroleum ether : diethyl ether under vacuum, 0.04% (v/v) Triton X-100 was added and the residue was resuspended into water. Reactions were performed in a 1:1 molar ratio of 9-*cis*-*β*-apo-10'-carotenal:AtCCD8.2 at a concentration of 12 µM in stopped flow assay buffer. For

determination of the kinetic parameters k_1 and k_2, the enzyme was diluted relative to the substrate.

pH rate profile experiments were performed using the continuous assay described in Section 8.9.7 with the exception that D27 buffer was replaced. For pH 4.12, 4.65, 5.13, 5.61 and 5.94 100 mM sodium acetate buffer was used, for pH 6.71 100 mM PBS buffer was used and for pH 7.70, 8.19, 8.62, 9.0, 9.62 and 10.3 100 mM tris-HCl was used.

Assays with group specific reagents were performed using the continuous assay method described in Section 8.9.7. AtCCD8 was preincubated with group specific reagents (Chapter 6.4) at a final concentration of 1 mM and pre-incubated with the enzyme for 10 minutes before the addition of substrate and monitoring at 430 nm.

Chapter Nine: Bibliography

1. S. Abe, A. Sado, K. Tanaka, T. Kisugi, K. Asami, S. Ota, H. I. Kim, K. Yoneyama, X. Xie, T. Ohnishi, Y. Seto, S. Yamaguchi, K. Akiyama, K. Yoneyama & T. Nomura, *Proc. Natl. Acad. Sci. USA*, (2014), doi: 10.1073/pnas.1410801111.
2. K. Akiyama, K. Matsuzaki & H. Hayahi, *Nature*, (2005), **435**, 824.
3. A. Alder, I. Holderman, P. Beyer & S. Al-Babili, *Biochem. J.*, (2008), **416**, 289.
4. A. Alder, M. Jamil, M. Marzorati, M. Bruno, M. Vermathen, P. Bigler, S. Ghisla, H. Bouwmeester, P. Beyer & S. Al-Babili, *Science*, (2012) **335**, 1348.
5. P. Auffinger, F. Hays, E. Westhof & P. Ho, *Proc. Natl. Acad. Sci. USA*, (2004), **101**, 16789.
6. M. E. Auldridge, A. Block, J. T. Vogel, C. Dabney-Smith, I. Mila, M. Bouzayen, M. Magallanes-Lundback, D. DellaPenna, D. R. McCarty & H. J. Klee, *Plant J.*, (2006a), **45**, 982.
7. M. E. Auldridge, D. R. McCarty & H. J. Klee, *Curr. Opin. Plant Biol.*, (2006b), **9**, 315.
8. P. G. Bartels & C. W. Watson, *Weed Sci.*, (1978), **26**, 198.
9. A. Beer, *Ann. Phys.*, (1852), **162**, 78.
10. S. W. Benson, K. W. Egger & D. M. Golden, *J. A. Chem. Soc.*, (1965), **87**, 468.
11. P. H. Bessette, F. Aslund, J. Beckwith & G. Georgiou, *Proc. Natl. Acad. Sci. USA*, (1999), **96**, 13703.
12. C. A. Beveridge, *Plant Growth Regul.*, (2000), **32**, 193.
13. C. A. Beveridge, J. J. Ross & I. C. Murfet, *Plant Physiol.*, (1996), **110**, 859.
14. J. Booker, M. Auldridge, S. Wills. D. McCarty, H. Klee & O. Leyser, *Curr. Biol.*, (2004), **14**, 1232.
15. J. Booker, T. Sieberer, W. Wright, L. Williamson, B. Willet, P. Stimberg, C. Turnbull, M. Srinivasan, P. Goddard & O. Leyser, *Dev. Cell*, (2005), **8**, 443.

16. T. Borowski, M. R. A. Blomberg & P. E. M. Siegbahn, *Chem. Eur. J.*, (2008), **14**, 2264.

17. F. Bouvier, C. Suire, J. Mutterer & B. Camara, *The Plant Cell*, (2003), **15**, 47.

18. F. Bouvier, J. C. Isner, O. Dogbo & B. Camara, *Trend. Plant. Sci.*, (2005), **10**, 187.

19. J. Boyd, Y. Gai, K. M. Nelson, E. Lukiwski, J. Talbot, M. K. Loewen, S. Owen, L. I. Zaharia, A. J. Cutler, S. R. Abrams & M. C. Loewen, *Bioorg. Med. Chem.*, (2009) **17**, 2902.

20. M. M. Bradford, *Anal. Biochem.*, (1976), **72**, 248.

21. P. M. Bramley. *Inhibition of carotenoid biosynthesis*, In A. J. Young & G. Britton, editors. *Carotenoids in Photosynthesis*, Springer; 1993.

22. P. B. Brewer, H. Koltai & C. A. Beveridge, *Mol. Plant.*, (2013), **6**, 18.

23. G. Britton, *Methods Enzymol.*, (1985), **111**, 113.

24. G. Britton, S. Ciaaen-Jensen, H. Pfander. Chapter 10: *Functions of Intact Carotenoids*, In G. Britton, S. Ciaaen-Jensen & H. Pfander, editors. *Carotenoids*, Volume 4: Natural Functions. Basel: Birkhauser Verlag; 2008.

25. M. Bruno, M. Hofmann, M. Vermathen, A. Alder, P. Beyer & S. Al-Babili, *FEBS Lett.*, (2014), **588**, 1802.

26. T. D. H. Bugg, *Tetrahedron*, (2003), **59**, 7075.

27. T. D. H. Bugg, *Biochem. Biophys. Acta*, (1993), **1202**, 258.

28. T. D. H. Bugg & S. Ramaswamy, *Curr Opin Chem Bio.*, (2003), **12**, 134.

29. T. D. H Bugg, M. Ahmad, E. M. Hardiman & R. Rahmanpour, *Nat. Prod. Rep.*, (2011), **28**, 1883.

30. T. D. H. Bugg, (2012), *Introduction to Enzyme and Coenzyme Chemistry*, 3rd Ed. Wiley-Blackwell, New Jersey.

31. A. Burbidge T. Grieve, A. Jackson, A. J. Thompson & I. Taylor, *J. Exp. Bot.*, (1997), **48**, 2111.

32. A. Burbidge, T. M. Grieve, A. Jackson, A. J. Thompson, D. R. McCarty & I. B. Taylor, *Plant J.*, (1999), **17**, 427.

33. N. A. Burgess-Brown, S. Sharma, F. Sobott, C. Loenarz, U. Oppermann & O. Gileadi, *Protein Expres. Purif.*, (2008), **59**, 94.

34. G.W. Burton, *J. Nutr.*, (1989), **119**, 109.

35. R. Campbell, L. J. M. Ducreux, W. L. Morris, J. A. Morris, J. C. Suttle, G. Ramsay, G. J. Bryan, P. E. Hedley & M. A. Taylor, *Plant Physiol.*, (2010), **154**, 656.

36. P. Chander, S. Gentleman, E. Poliakov & T. M. Redmond, *J. Biol. Chem.*, (2012), **287**, 30552.

37. J. Chandler, PhD thesis, (2014), University of Warwick & Syngenta.

38. W.-H. Cheng, A. Endo, L. Zhou, J. Penney, H. Chen, A. Arroyo, P. Leon, E. Nambara, T. Asami, M. Seo, T. Koshiba & J. Sheen, *Plant Cell*, (2002) **14**, 2723.

39. L. Cholnoky, K. Gyorgyfy, A. Ronai, J. Szabolcs, G. Toth, G. Galasko, A. K. Mallams, E. S. Waight & B. C. L. Weedon, *J. Chem. Soc. C*, (1969), 1256.

40. D. W. Christianson & W. N. Lipscomb, *Acc. Chem. Res.*, (1989), **22**, 62.

41. R. J. Cogdell & H. A. Frank, *Biochim. Biophys. Acta*, (1987), **895**, 63.

42. R. Creelman, E. Bell & J. Mullet, *Plant Physiol.*, (1992), **99**, 1258.

43. P. S. Deigner, W. C. Law, F. J. Canada & R. R. Rando, *Science*, (1889), **244**, 968.

44. K. P. Dowd, R. Hershline, S. W. Ham & S. Naganathan, *Nat. Prod. Rep.*, (1994), **11**, 251.

45. E. A. Dun, J. B. Ferguson & C. A. Beveridge, *Plant Physiol.*, (2006), **142**, 812.

46. C. Emenhiser, L. C. Sander & S. J. Schwartz, *J. Chromatogr. A*, (1995), **707**, 205.

47. J. A. Fee, K. L. Findling, T. Yoshida, R. Hille, G. E. Tarr, D. O. Hearshen, W. R. Dunham, E. P. Day, T. A. Kent & E. Munck, *J. Biol. Chem.*, (1983), **259**, 124.

48. D. J. Ferraro, L. Gakhar & S. Ramaswamy, *Biochem. Biophys. Res. Comm.*, (2005), **338**, 175.

49. A. Fielding, E. Kovaleva, E. Farquhar, J. Lipscomb & L. Que Jr., *J. Bio. Inorg. Chem.*, (2011), **16**, 341.

50. W. E. Finch-Savage & G. Leubner-Metzger, *New Phytol.*, (2006), **171**, 501.

51. D. S. Floss, W. Schliemann, J. Schmidt & D. Strack, *Plant Physiol.*, (2008), **148**, 1267.

52. C. S. Foote & R. W. Penny, *J. Am. Chem. Soc.*, (1968) **90**, 6233.

53. H. A. Frank & R. L. Christensen, *Excited Electronic Sates, Photochemistry and Photophysics of Carotenoids*, In *Carotenoids*, G. Britton, H. Pfander & H. C. S. Liaan-Jensen (Eds.), (2008), **4**, 167.

54. N.P. Franks & P. Brick, *Structure*, (1996), **4**, 287.

55. R. R. French, P. Holzer, M. G. Leuenberger & W. D. Woggon, *Angew. Chem. Intl. Ed.*, (2000), **39**, 1267.

56. W. Fu, J. Lin & P. Cen, *Appl. Microbiol. Biotchnol.*, (2007), **75**, 777.

57. C. Garcia-Limones, K. Schnabele, R. Blanco-Portales, M. L. Bellido, J. L. Caballero, W. Schwab & J. Munoz-Blanco, *J. Agric. Food. Chem.*, (2008), **56**, 9277.

58. A. Gibello, E. Ferrer, M. Martin & A. Garrido-Pertierra, *Biochem. J.*, (1994) **301**, 145.

59. S. A. Gillmor, A. Villasenor, R. Fletterick, E. Sigal & M. F. Browner, *Nat. Struc. Mol. Biol.*, (1997), **4**, 1003.

60. V. Gomez-Roldan, S. Fermas, P. B. Brewer, V. Puech-Pages, E. A. Dun, J.-P. Pillot, F. Letisse, R. Matusova, S. Danoun, J.-C. Portais, H. Bouwmeester, G. Becard, C. A. Beveridge, C. Rameau & S. F. Rochange, Nature, (2008), **455**, 189.

61. M. González-Guzmán, N. Apostolova, J. Belles, J. Barrero, P. Piqueras, M. Ponce, J. Micol, R. Serrano & P. Rodriquez, *Plant Cell*, (2002), **14**, 1833.

62. D. S. Goodman & H. S. Huang, *Science*, (1965), **149**, 879.

63. M. Goujon, H. McWilliam, W. Li, F. Valentin, S. Squizzato, J. Paern & R. Lopez, *Nucl. Acids Res.*, 2010, **38**, (suppl. 2), W695.

64. A. L. Green, G. L. Sainsbury, B. Saville & M. Stansfield, *J. Chem. Soc.*, (1958), **316**, 1583.

65. S.-Y. Han, H. Inoue, T. Terada, S. Kamoda, Y. Saburi, K. Sekimata, T. Saito, M. Kobayashi, K. Shinozaki, S. Yoshida & T. Asami, *Bioorg. Med. Chem. Lett.*, (2002), **12**, 1139.

66. S.-Y. Han, H. Inoue, T. Terada, S. Kamoda, Y. Saburi, K. Sekimata, T. Saito, M. Kobayashi, K. Shinozaki & S. Yoshida, *J. Enzyme Inhib. Med. Chem.*, (2003), **18**, 279.

67. S.-Y. Han, N. Kitahata, K. Sekimata, T. Saito, M. Kobayashi, K. Nakashima, K. Yamaguchi-Shinozaki, S. Yoshida & T. Asami, *Plant Physiol.*, (2004a), **135,** 1574.

68. S-Y. Han, N. Kitahata, T. Saito, M. Kobayashi, K. Shinozaki, S. Yoshida & T. Asami, *Bioorg. Med. Chem. Lett.*, (2004b), **12**, 3033.

69. P. J. Harrison & T. D. H. Bugg, *Arch. Biochem. Biophys.*, (2014), **544**, 105.

70. F.-C. Huang, G. Horváth, P. Molnár, E. Turcsi, J. Deli, J. Schrader, G. Sandmann, H. Schmidt & W. Schwab, *Phytochemistry*, (2009), **70**, 457.

71. F.-C. Huang, P. Molnar & W. Schwab, *J. Exp. Bot*., (2009), **60**, 3011.

72. H.-K. Hund, J. Breuer, F. Lingens, J. Hüttemann, R. Kappl & S. Fetzneral, *Eur. J. Biochem.*, (1999), **263**, 871.

73. M. Ibdah, Y. Azulay, V. Portnoy, B. Wasserman, E. Bar, A. Meir, Y. Burger, J. Hirschberg, A. Schaffer, N. Katzir, Y. Tadmor, & E. Lewinsohn, *Phytochemistry*, (2006), **67**, 1579.

74. A. Ilg, P. Beyer, S. Al-Babili, *FEBS J.*, (2009), **276**, 736.

75. H. Inoue, H. Nojima & H. Okayama, *Gene*, (1990), **96**, 23.

76. T. Isaacson, G. Ronen, D. Zamir & J. Hirschberg, *Plant Cell*, (2002), **14**, 333.

77. T. Isaacson, I. Ohad, P. Beyer & J. Hirschberg, *Plant Physiol.*, (2004), **136**, 4246.

78. S. Iuchi, M. Kobayashi, T. Taji, M. Naramoto, M. Seki, T. Kato, S. Tabata, Y. Kakubari, K. Yamaguchi-Shinozaki & K. Shinozaki, *Plant J.*, (2001), **27**, 325.

79. N. Jin, Q. Yuan, S. Li & G. H. Travis, *J. Biol. Chem.*, (2007), **282**, 20915.

80. S. Kamoda & Y. Saburi, Bioscience, *Biotechnol. Biochem.*, (1993), **57**, 926.

81. J. F. Kane, *Curr. Opin. Biotechnol.*, (1995), **6**, 494.

82. P. Karrer, A. Helfenstein, H. Wehrli & A. Wettstein, *Helv. Chim. Acta*, (1930), **13**, 1084.

83. M. Kato, H. Matsumoto, Y. Ikoma, H. Okuda & M. Yano, *J. Exp. Bot.*, (2006), **57**, 2153.

84. C. Kiefer, S. Hessel, J.M. Lampert K. Vogt, M.O. Lederer, D.E. Breithaupt & J. Von Lintig, *J. Biol. Chem*., (2001), **276**, 14110.

85. Y.-S. Kim, E.-S. Seo & D.-K. Oh, *Biotechnol. Lett*., (2012), **34**, 1851.

86. P. D. Kiser, M. Golczak, D. T. Lodowski, M. R. Chance & K. Palczewski, *Proc. Natl. Acad. Sci. USA,* (2009), **106**, 17325.

87. P. D. Kiser, E. Farquhar, W. Shi, X. Sui, M. R. Chance & K. Palczewski, *Proc. Natl. Acad. Sci. USA*, (2012), **109**, E2747.

88. N. Kitahata, S-Y. Han, N. Noji, T. Saito, M. Kobayashi, T. Nakano, K. Kuchitsu, K. Shinozaki, S. Yoshida, S. Matsumoto, M. Tsujimoto & T. Asami, *Bioorg. Med. Chem.*, (2006), **14**, 5555.

89. D. P. Kloer, S. Ruch, S. Al-Babili, P. Beyer & G. E. Schulz, *Science*, (2005), **308**, 267.

90. G. Kudla, A. W. Murray, D. Tollervey & J. B. Plotkin, *Science*, (2009), **324**, 255.

91. B. Kucera, M. A. Cohn & G. Leuber-Metzger, *Seed Sci. Res.*, 2005, **15**, 281.

92. J. M. Lampert, J. Holzschuh, S. Hessel, W. Driever, K. Vogt & J. Von-Lintig, *Development*, (2003), **130**, 2173

93. C. R. D. Lancaster, A. Kröger, M. Auer & H. Michel, *Nature*, (1999), **402**, 377.

94. G. Lazar & H. M. Goodman, *Proc. Natl. Acad. Sci. USA*, (2006), **103**, 472.

95. V. Lefebvre, H. North, A. Frey, B. Sotta, M. Seo, M. Okamoto, E. Nambara & A. Marion-Poll, *Plant J.*, (2006), **45**, 309.

96. M. G. Leuenberger, C. Engeloch-Jarret & W-D. Woggon, *Ang. Chem. Int. Ed.*, (2001), 40, 2613.

97. O. Leyser, *Curr. Opin. Genet. Dev.*, (2005), **15**, 468.

98. H. Lin, R. Wang, Q. Qian, M. Yan, X. Meng, Z. Fun, C. Yan, B. Jiang, Z. Su, J. Li & Y. Wang, *Plant Cell*, (2009), 21, 1512.

99. E. K. Marasco & C. Schmidt-Dannert, *ChemBioChem*, (2008), **9**, 1450.

100. S. Mathieu, N. Terrier, J. Procureur, F. Bigey & Z. Gunata, *J. Exp. Bot.*, (2005), **56**, 2721.

101. A. de Marco, *Nat. Protoc.*, (2007), **2**, 2632.

102. J. K. McBee, V. Kuksa. R. Alvarez, A. R. de Lera, O. Prezhdo, F. Haeseleer, I. Sokal & K. Palczewski, *Biochemistry*, (2000), **39**, 11370.

103. F. McCapra, *Acc. Chem. Res.*, (1976), **9**, 201.

104. R. Mendelsohn, *Climatic Change* (2007), 81, **1**, 61.

105. Merck Millipore Competent Cells selection guide. http://www.emdmillipore.com/INTERSHOP/web/WFS/Merck-GB-Site/en_US/-/GBP/ShowDocument-Pronet?id=201306.9863. Accessed on 20/10/2014 at 15.33 PM.

106. S. Messing, S. Gabelli, I. Echeverria, J. T. Vogel, J.-C. Guan, B.-C. Tan, H. J. Klee, D. R. McCarty & L.-M. Amzel, *Plant Cell*, (2010), **22**, 2970.

107. A. Mihalyi, PhD thesis, (2014), University of Warwick.

108. G. Moiseyev, Y. Chen, Y. Takahashi, B. X. Wu & J. Ma, *Proc. Natl. Acad. Sci. USA*, (2005), **102**, 12413.

109. G. Moiseyev, Y. Takahashi, Y. Chen, S. Gentleman, T. M. Redmond, R. K. Crouch & J. Ma, *J. Biol. Chem.*, (2006), **281**, 2835.

110. T. Moore, *Biochem. J.*, (1930), **24**, 692.

111. S. E. Morris, C. G. Turnball, I. C. Murfet & C. A. Beveridge, *Plant Physiol.*, (2001), **126**, 1205.

112. E. Nakamaru-Ogiso, T. Yano, T. Ohnishi & T. Yagi, *J. Biol. Chem.*, (2001), **277**, 1608.

113. E. Nambara & A. Marion-Poll, *Annu. Rev. Plant Biol.*, (2005), **56**, 165.

114. S. J. Neill, E. C. Burnett, R. Desikan & J. T. Hancock, *J. Exp. Bot.*, (1998), **49**, 1893.

115. D. M. Niedzwiedzki, D. J. Sandberg, H. Cong, M. N. Sandberg, G. N. Gibson, R. R. Birge & H. A. Frank, *Chem. Phys.*, (2009), **357**, 4.

116. C. Notredame, D. G. Higgins & J. Heringa, *J. Mol. Biol.*, (2000), **302**, 205.

117. V. Oberhauser, O. Voolstra, A. Bangert, J. Von Lintig & K. Vogt, *Proc. Natl. Acad. Sci. USA*, (2008), **105**, 19000.

118. D. H. Ohlendorf, J. D. Lipscomb & P. C. Weber, *Nature*, (1988), **336**, 403.

119. A. Ohmiya, S. Kishimoto, R. Aida, S. Yoshioka & K. Sumitomo, *Plant Physiol.*, (2006), **142**, 1193.

120. T. Oka & F. J. Simpson, *Biochem. Biophys. Res. Comm.*, (1971), **43**, 1.

121. J. A. Olson & O. Hayaishi, *Proc. Natl. Acad. Sci. USA*, (1965), **54**, 1364.

122. W. H. Ohme-Johnson, *Ann. Rev. Biochem.*, (1973), **42**, 159.

123. W. H. Orme-Johnson & N. R. Orme-Johnson, *Methods Enzymol.*, (1978), **53**, 259.

124. G. Palmer & R. H. Sands, *J. Biol. Chem.*, (1966), **241**, 253.

125. H. Park, S. S. Kreunen, A. J. Cuttriss, D. DellaPenna & B. J. Pogson, *Plant Cell*, **14**, 321.

126. A. Prado-Cabrero, D. Scherzinger, J. Avalos & S. Al-Babili, *Eukaryotic Cell*, (2007), **6**, 650.

127. Promega NAD(P)H Glo Kit product manual. http://www.promega.co.uk/~/media/files/resources/protocols/technical%20manuals/101/nadp%20nadph%20glo%20assay%20protocol.pdf. Accessed on 20/09/2014 at 12.30 PM.

128. L. Que Jr., J. Widom & R. L. Crawford, *J. Biol. Chem.*, (1981), **256**, 10941.

129. T. M. Redmond, S. Gentleman, T. Duncan, S. Yu, B. Wiggert, E. Gnatt & F.X. Cunningham, *J. Biol. Chem.*, (2001), **276**, 6560.

130. T. M. Redmond, E. Poliakov, S. Kuo, P. Chander & S. Gentleman, *J. Biol. Chem.*, (2010), **285**, 1919.

131. S. Rottem & O. Markowitz, *J. Bacteriol.*, (1979), **140**, 944.

132. A. Rubio, J.L. Rambla, M. Santaella, M.D. Gómez, D. Orzaez, A. Granell & L. Gómez- Gómez, *J. Biol. Chem.*, (2008), **283**, 24816.

133. S. Ruch, P. Beyer, H. Ernst & S. Al-Babili, *Mol. Microbiol*., (2005) **55**, 1015.

134. P. Sainsbury, PhD thesis, (2014), University of Warwick.

135. J. Sambrook & D.W. Russell, *Molecular Cloning A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 2001, New York, U.S.

136. J. Sanvoisin, G. J. Langley & T. D. H. Bugg, *J. Am. Chem. Soc*., (1995), **117**, 7836.

137. H. M. Schaefer, V. Schaefer & D. J. Levey, *Trends Ecol. Evol.*, (2004), **19**, 577.

138. D. Scherzinger, E. Scheffer, C. Bär, H. Ernst & S. Al-Babili, *FEBS J*., (2010), **277**, 4662.

139. H. Schmidt, R. Kurtzer, W. Eisenreich & W. Schwab, *J. Biol. Chem.*, (2006), **281**, 9845.

140. S. H. Schwartz, X. Qin & J. A. D. Zeevaart, *J. Biol. Chem.*, (2001) **276**, 25208.

141. S. H. Schwartz, B.-C. Tan, D. Gage & J. A. D. Zeevaart, D. R. McCarty, *Science*, (1997), **275**, 1872.

142. S. H. Schwartz, B.-C. Tan, D. R. McCarty, W. Welch & J. A. D. Zeevaart, *Biochim. Biophys. Acta*, (2003), **1619**, 9.

143. S. H. Schwartz, X. Qin & M. C. Loewen, *J. Biol. Chem.*, (2004), **279**, 46940.

144. M. Seo, H. Koiwai, S. Akaba, T. Komano, T. Oritani, Y. Kamiya & T. Koshiba, *Plant J*., (2000), **23**, 481.

145. M. J. Sergeant, P. J. Harrison, R. Jenkins, G. Moran, T. D. H. Bugg & A. J. Thompson, *New J. Chem.*, (2013), **37**, 3461.

146. M. J. Sergeant, J.J.-Li, C. Fox, N. Brookbank, D. Rea, T. D. H. Bugg & A. J. Thompson, *J. Biol. Chem.*, (2009), **284**, 5257.

147. M. J. Sergeant, J. Chandler, P. J. Harrison, J.-J. Li, T. D. H. Bugg & A. J. Thompson, unpublished.

148. K. E. Sharpless, J. B. Thomas, L. C. Sander & S. A. Wise, *J. Chromatogr. B*, (1996), **678**, 187.

149. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Sodling, J. D. Thompson & D. Higgins, *Mol. Syst. Biol.*, (2011), **7**, 539.

150. A. J. Simkin, S. Schwartz, M. Auldridge, M. Taylor & H. J. Klee, *Plant J.*, (2004), **40**, 882.

151. A. J. Simkin, H. Moreau, M. Kuntz, G. Pagny, C. Lin, S. Tanksley & J. McCarthy, *J. Plant Physiol.*, (2008), **165**, 1087.

152. P. Singh, L. Sharma, S. R. Kulothungan, B. V. Adkar, R. S. Prajapati, P. S. S. Ali, B. Krishnan & R. Varadarajan, *PLoS One*, (2013), **8**, e63442.

153. K. Sorefan, J. Booker, K. Haurogne, M. Goussot, K. Bainbridge, E. Foo, S. Chatfield, S. Ward, C. Beveridge, C. Rameau & O. Leyser, *Genes and Dev.*, (2003), **17**, 1469.

154. E. L. Spence, G. J. Langley & T. D. H. Bugg, *J. Am. Chem. Soc.*, (1996), **118**, 8336.

155. C. G. Spilianakis, G. R. Lee & R. A. Flavell, *Eur. J. Immunol.*, (2005), **35**, 3400.

156. J. Stubbe & W. van der Donk, *Chem. Rev.*, (1998), **98**, 705.

157. X. Sui, P. D. Kiser, J. von-Lintig & K. Palczewski, *Arch. Biochem. Biophys.*, (2013), **539**, 203.

158. X. Sui, P. D. Kiser, T. Che, P. R. Carey, M. Golczak, W. Shi, J. von-Lintig & K. Palczewski, *J. Biol. Chem.*, (2014), **289**, 12286.

159. J. W. Suttie, *Annu. Rev. Biochem.*, (1985), **54**, 459.

160. K. Tagawa & D. I. Arnon, *Biochem. Biophys. Acta*, (1968), **153**, 602.

161. B.-C. Tan, S. H. Schwartz, J. A. D. Zeevaart & D. R. McCarty, *Proc. Natl. Acad. Sci. USA*, (1997), **94**, 12235.

162. B.-C. Tan, K. Cline & D. R. McCarty, *Plant J.*, (2001), **27**, 373.

163. B.-C. Tan, L. M. Joseph, W.-T. Deng, L. Liu, Q.-B. Li, K. Cline & D. R. McCarty, *Plant J.*, (2003), **35**, 44.

164. I. B. Taylor, T. Sonneveld, T. D. H. Bugg & A. J. Thompson, *J. Plant Growth Regul.*, (2005), **24**, 253.

165. L. H. Tonucci, J. M. Holden, G. R. Beecher, F. Khachik, C. S. Davis & G. Mulokozi, *J. Agric. Food Chem.*, (1995), **43**, 579.

166. K. Tsukida & K. Saiki, *J. Chromatogr. A*, (1982), **3**, 359.

167. T. Tuller, Y. Y. Waldman, M. Kupiec & E. Ruppin, *Proc. Natl. Acad. Sci. USA*, (2010), **107**, 3645.

168. C. G. Turnbull, J. P. Booker & O. Leyser, *Plant J.*, (2002), **32**, 255.

169. M. Umehara, A. Hanada, S. Yoshidam K. Akiyama, T. Arite, N. Takeda-Kamiya, H. Magome, Y. Kamiya, K. Shirasu, K. Yoneyama, J. Kyozuka & S. Yamaguchi, *Nature*, (2008), **455**, 195.

170. United Nations Department of Economic and Social Affairs: Population Division, Population Estimates and Projections Section. World Population Estimates, 2012 revision. http://esa.un.org/unpd/wpp/unpp/panel_population.htm. Accessed on 14/07/2014 at 17.07 PM.

171. W. A. Van Der Donk, A-L. Tsai & R. J. Kulmacz, *Biochemistry*, (2002), **41**, 15451.

172. J. Van Norman, J. Zhang, C. Cazzonelli, B. Pogson, P. J. Harrison, T. D. H. Bugg, K.-X. Chan, A. J. Thompson & P. N. Benfrey, *Proc. Natl. Acad. Sci. USA*, (2014), **111**, E1300.

173. J. T. Vogel, B.-C. Tan, D. R. McCarty & H. J. Klee, *J. Biol. Chem.*, (2008), **283**, 11364.

174. J. Von-Lintig & K. Vogt, *J. Biol. Chem.*, (2000), **275**, 11915.

175. G. Wald, *Nature*, (1968), **219**, 800.

176. M. H. Walter & D. Strack, *Nat. Prod. Rep.*, (2011), **28**, 663.

177. M. H. Walter, D. Floss & D. Strack, *Planta*, (2010), **232**, 1

178. M. T. Waters, P. B. Brewer, J. D. Bussell, S. M. Smith & C. A. Beveridge, *Plant Physiol.*, (2012), **159**, 1073.

179. A. K. Whiting, Y. R. Boldt, M. P. Hendrich, L. P. Wackett & L. Que Jr., *Biochemistry*, (1996), **35**, 160.

180. D. H. Williams, I. Fleming, authors. Chapter 1: Ultraviolet and Visible Spectra, In D. H. Williams, I. Fleming, authors. *Spectroscopic Methods in Organic Chemistry*, Fifth Edition. London: McGraw Hill; 1995.

181. P. Winterhalter & M. Straubinger, *Food Rev. Int.*, (2000), **16**, 39.

182. P. Winterhalter, R. Rouseff, editors. Chapter 1: *An Introduction,* In P. Winterhalter, R. Rouseff, authors. *Carotenoid Derived Aroma Compounds*. Washington: ACS Symposium Series; 2001.

183. G. M. Wirtz, C. Bornemann, A. Giger, R. K. Muller, H. Schneider, G. Schlotterbeck, G. Schiefer & W. D. Woggon, *Helv. Chim. Acta.*, (2001), **84**, 2301.

184. A. Wyss, G. Wirtz, W. Woggon, R. Brugger, M. Wyss, A. Friedlein, H. Bachmann & W. Hunziker, *Biochem. Biophys. Res. Commun.*, (2000), **271**, 334.

185. F. F. Xu & J. A. Imlay, *Appl. Environ. Microbiol.*, (2012), **78**, 3614.

186. W. Yan, G-F. Jang, F. Haeseleer, N. Esumi, J. Chang, M. Kerrigan, M. Campochiaro, P. Campochiaro, K. Palczewski & D. J. Zack, *Genomics*, (2001), **72**, 193.

187. Y. Yu, D. Kalinowski, A. Siafakas, P. Jansson, C. Stefani, D. B. Lovejoy, P. C. Sharpe, P. V. Bernhardt & D. R. Richardson, *J. Med. Chem.*, (2009), **52**, 5271.

188. J. A. D. Zeevaart, T. G. Heath & D. A. Gage, *Plant Physiol.*, (1989), **91**, 1594.

189. Y. Zhang, A. D. J. van Dijk, A. Scaffidi, G. R. Flematti, M. Hofmann, T. Charnikhova, F. Verstappen, J. Hepworth, S. van der Krol, O. Leyser, S. M. Smith, B. Zwanenburg,

S. Al-Babili, C. Ruyter-Spira & H. J. Bouwmeester, *Nature Chem. Biol.*, (2014), doi: 10.1038/NCHEMBIO.1660.

Chapter Ten: Appendices

10.1 Full CCD Sequence Alignment

Full T-Coffee sequence alignment of selected members of the carotenoid cleavage dioxygenase enzyme family and other related proteins. Conserved histidine residues are marked in red, second shell glutamic acid residues in blue and conflicts in orange. Bt – *Bos taurus*; Hs – *Homo sapiens*; Zm - *Zea mays*; At – *Arabidopsis thaliana*; S. – *Synechocystis sp.*; Dm – *Drosophilia melanogaster*; Le – *Solanum lycopersicum*; Ps – *Pisum sativum*; Pa – *Persea americana*; Bo – *Bixa orellana*; Cs – *Crocus sativum*; Mt – *Mycobacterium tuberculosis*; Mm – *Mus musculus*; Na – *Novosphingobium aromaticivorans*; Pp – *Pseudomonas paucimobilis*; Gf - *Gibberella fujikuroi*.

```
AN_Q28175|_BtRPE65    1 MSSQVEH-------------------------------------------------------    7
AN_Q16518|_HsRPE65    1 MSIQVEH-------------------------------------------------------    7
AN_Q9BYV7|_HsBCO2     1 MFFRVFL-HFI---------------------------------------------------   10
AN_Q9HAY6|_HsBCO1     1 MDIIFG--------------------------------------------------------    6
AN_O24592|_ZmNCED1    1 MQGLAPP-TSV---------------------------------------------------   10
AN_Q45VT7|_ZmCCD1     1 MGTEAEQ-PDMD--------------------------------S---------------   12
AN_Q9M9F5|_AtNCED9    1 MTIITII-SGMYIYSLLSQDAHHSQYGQNTNLVLKKPIPKPQTAAFNQESTMASTTLLPSTST   62
AN_Q9LRR7|_AtNCED3    1 MASFTAT-AAV----------------------------------SG---------------   12
AN_Q9C6Z1|_AtNCED5    1 MACSYILTPNP---------------------------------------------------   11
AN_O49505|_AtNCED2    1 MVSLLTM-PMS---------------------------------------------------   10
AN_Q9LRM7|_AtNCED6    1 MQHSLRS-DLLP--------------------------------TK---------------   13
AN_O49675|_AtCCD4     1 MDSVSSS-SFLS--------------------------------S---------------   12
AN_O65572|_AtCCD1     1 MAEKLSD-G-----------------------------------------------------    8
AN_Q7XJM2|_AtCCD7     1 MSLPIPP-KFLP--------------------------------P---------------   12
AN_Q8VY26|_AtCCD8     1 MASLITT-KAM---------------------------------------------------   10
AN_P74334|_S.ACO      1 MVTSPPT-SS----------------------------------------------------    9
AN_Q9NKW9|_DmNinaB    1 MAAGVFK-SFM---------------------------------------------------   10
AN_Q6E4P5|_LeCCD1a    1 MGRKEDD-GV----------------------------------------------------    9
An_Q8LP16|_PsNCED2    1 MATFTAS-PSN---------------------------------------------------   10
AN_Q8LP15|_PsNCED3    1 MAPSSLALNSN---------------------------------------------------   11
AN_Q9AXZ3|_PaNCED1    1 MTTIRQK-PKT--------------------------------F----------------   11
AN_Q9AXZ4|_PaNCED3    1 MSMATPT-TTC--------------------------------GA---------------   12
AN_C5H805|_LeNCED2    1 -------------------------------------------------------------    0
AN_Q70YP8|_BoLCO      1 MQ------------------------------------------------------------    2
AN_Q84K96|_CsZCD      1 MQ------------------------------------------------------------    2
AN_Q8LP17|_PsCCD1     1 MGSEKKE-N-----------------------------------------------------    8
AN_O06785|_MtCCO      1 MTTAQAA-E-----------------------------------------------------    8
AN_Q99NF1|_MmBCO2     1 MLGPKQS-------------------------------------------------------    7
AN_Q9JJS6|_MmBCO1     1 MEIIFG--------------------------------------------------------    6
AN_Q2GA76|_NaCCO      1 MAQFPNT-P-----------------------------------------------------    8
AN_Q53353|_PpLSD      1 MAHFPQT-P-----------------------------------------------------    8
An_O24023|_LeNCED1    1 MATTTSH-ATN---------------------------------------------------   10
AN_Q5GN50|_GfCarX     1 MKFLQQN-SFTQ--------------------------------T---------------   12
cons                  1                                                               63

AN_Q28175|_BtRPE65    8 -------------------------------------------------------------    7
AN_Q16518|_HsRPE65    8 -------------------------------------------------------------    7
AN_Q9BYV7|_HsBCO2    11 -------------------------------------------------------------   10
AN_Q9HAY6|_HsBCO1     7 -------------------------------------------------------------    6
AN_O24592|_ZmNCED1   11 -S-I--HR----HL-PAR----SR---ARASNS------VRFSP----------RAVSSVPPA   41
```

```
AN_Q45VT7|_ZmCCD1    13 ------------------------------------------------------------   12
AN_Q9M9F5|_AtNCED9   63 QF-L--DRTF--S---T----SSS--SSRPKLQSLSFSSTLRNKKLVVPCYVSSSVNKKSSVS  111
AN_Q9LRR7|_AtNCED3   13 RW-L--GGNH--TQPPL----SSSQSSDLSYCS-----SLPMAS----------RVTRKLNVS   51
AN_Q9C6Z1|_AtNCED5   12 TK-L--NL---------SFAPSDLD-APSPSSS-VS---FTNTK----------PRRRKLSAN   47
AN_O49505|_AtNCED2   11 GG-I--KT--------------------WPQAQ-ID---LGFRP----------IKRQPKVIK   36
AN_Q9LRM7|_AtNCED6   14 TS-P--RS-H--LL-PQ------------P-----------KNA----------NISRRILIN   36
AN_O49675|_AtCCD4    13 TFSL--HH----S-------------------------LLRRRS---------SSPTLLRIN   34
AN_O65572|_AtCCD1     9 ------------------------------------------------------------    8
AN_Q7XJM2|_AtCCD7    13 LKSPPIHHH-------------------------------------------------------   21
AN_Q8VY26|_AtCCD8    11 MS-H--HH----VL-S------ST----RIT--------TLYSD-----------------N   29
AN_P74334|_S.ACO     10 ----PS----------------------------------------------------------   11
AN_Q9NKW9|_DmNinaB   11 ------------------------------------------------------------   10
AN_Q6E4P5|_LeCCD1a   10 ------------------------------------------------------------    9
An_Q8LP16|_PsNCED2   11 TW-I--NTKS--SS--------------RTNSS-----SILLNKK---------RSTSKNTIS   40
AN_Q8LP15|_PsNCED3   12 TW-A--TTTNKPQF-PQSFSISSSS-STSFNDK-ST---LKLKK----------PNRKMLLLQ   55
AN_Q9AXZ3|_PaNCED1   12 ---------------TI----H-------------------------------------------   14
AN_Q9AXZ4|_PaNCED3   13 GDLL--QNPK--LL-PI----SKN--LSRPKNF-I---MLKHNTP----------LIQCCSHS   50
AN_C5H805|_LeNCED2    1 ------------------------------------------------------------    0
AN_Q70YP8|_BoLCO      3 ------------------------------------------------------------    2
AN_Q84K96|_CsZCD      3 ------------------------------------------------------------    2
AN_Q8LP17|_PsCCD1     9 ------------------------------------------------------------    8
AN_O06785|_MtCCO      9 ------------------------------------------------------------    8
AN_Q99NF1|_MmBCO2     8 ------------------------------------------------------------    7
AN_Q9JJS6|_MmBCO1     7 ------------------------------------------------------------    6
AN_Q2GA76|_NaCCO      9 ------------------------------------------------------------    8
AN_Q53353|_PpLSD      9 ------------------------------------------------------------    8
An_O24023|_LeNCED1   11 TW-I--KTKL--SM-PS----SKE--FGFASNS-I---SLLKNQ----------HNRQSLNIN   47
AN_Q5GN50|_GfCarX    13 SMSQPH--E-------------------------------------------------------   19
cons                 64                                                             126

AN_Q28175|_BtRPE65    8 ------------------------------------------------------------    7
AN_Q16518|_HsRPE65    8 ------------------------------------------------------------    7
AN_Q9BYV7|_HsBCO2    11 ------------------------------------------------------------   10
AN_Q9HAY6|_HsBCO1     7 ------------------------------------------------------------    6
AN_O24592|_ZmNCED1   42 ECL-QAPFHKPVADLPAPSRK---P----------AAIAVPGHAAAP---------RKAEGGK   81
AN_Q45VT7|_ZmCCD1    13 ----------------------------H-RNDGVVVV-----------------P---A   23
AN_Q9M9F5|_AtNCED9  112 SSL-QS---------PTF--K---PPS-WKKLCNDVTNLIPK---------------TTNQN  142
AN_Q9LRR7|_AtNCED3   52 SAL-HTP---PALHFPKQSSN---SP---------AIVVKPK--------------AKESNT   83
AN_Q9C6Z1|_AtNCED5   48 SVS-DTP---NLLNFPNY-------------P--SPNPII-----------------PEKDT   73
AN_O49505|_AtNCED2   37 CTV-QID----VTELTKK------RQL-FT--P--RTTATP----------------PQHNP   66
AN_Q9LRM7|_AtNCED6   37 P-F-------KIPTLPDLTS-----P---------VP--SP--------------VKLKPTY   60
AN_O49675|_AtCCD4    35 SAVVEER---SPITNPSDNNDRRNKPKTLHN-RTNHTLVS-----------------S---P   72
AN_O65572|_AtCCD1     9 ---------------------------------SSIIS-----------------V---H   15
AN_Q7XJM2|_AtCCD7    22 ----QTP---PPLAPPRAAIS----------------ISIPD----------------TGLG   44
AN_Q8VY26|_AtCCD8    30 S------------------IG-------------DQQIK-----------------T---K   39
AN_P74334|_S.ACO     12 -----QR---SYS----------------------------------------------   16
AN_Q9NKW9|_DmNinaB   11 ------------------------------------------------------------   10
AN_Q6E4P5|_LeCCD1a   10 ----------------------------E-RIEGGVVV-----------------V---N   20
An_Q8LP16|_PsNCED2   41 CSL-QT-----TLPFPKKYQP-QSTNT-------TTTTLIPTRETKPNLPS---NTKPLHKQE   86
AN_Q8LP15|_PsNCED3   56 CAV-HSP---SVLDYPKQSYK---QPL-INK-ENDTETIIH-----------------KPKES   92
AN_Q9AXZ3|_PaNCED1   15 SSL-HSS---PVLHLPKLLTT-TTTPL-HEK-SQRELG----------------LILQEPNR   53
AN_Q9AXZ4|_PaNCED3   51 PSS-SSA---AVLHLPPKQPT-KSKPS-IKK-GEKSSTLTPSIEKNPGSHQVKTDQSGPNRVG  106
AN_C5H805|_LeNCED2    1 ------------------------------------------------------------    0
AN_Q70YP8|_BoLCO      3 ------------------------------------------------------------    2
AN_Q84K96|_CsZCD      3 ------------------------------------------------------------    2
AN_Q8LP17|_PsCCD1     9 ---------------------------------GVILE-----------------V---E   15
AN_O06785|_MtCCO      9 ------------------------------------------------------------    8
AN_Q99NF1|_MmBCO2     8 ------------------------------------------------------------    7
AN_Q9JJS6|_MmBCO      7 ------------------------------------------------------------    6
AN_Q2GA76|_NaCCO      9 ------------------------------------------------------------    8
AN_Q53353|_PpLSD      9 ------------------------------------------------------------    8
An_O24023|_LeNCED1   48 SSL-QAP---PILHFPKQSSN-YQTPK-------NNTISHPKQENN----------NSSSSST   88
AN_Q5GN50|_GfCarX    20 ----DVS---PAI----------------------------------------------   25
cons                127                                                             189

AN_Q28175|_BtRPE65    8 ---------------------------------------------------PA-GGYKKLF   16
AN_Q16518|_HsRPE65    8 ---------------------------------------------------PA-GGYKKLF   16
AN_Q9BYV7|_HsBCO2    11 ------RSHSATAVDFLPVMVH-RL--PVFKRYMG-NTPQKKAVFGQCR----GL-PCVAPLL   58
AN_Q9HAY6|_HsBCO1     7 ------------------------------------------------------------R-    7
```

```
AN_O24592|_ZmNCED1   82 KQLNLFQRAAAAALDAFEEGFV-ANVLE------R-PHGLPS------TA--DPA-VQIAGNF  127
AN_Q45VT7|_ZmCCD1    24 PR--PRKGLASWALDLLESLAV-RL--------------GHD------KT--KPL-HWLSGNF   60
AN_Q9M9F5|_AtNCED9  143 PKLNPVQRTAAMVLDAVENAMI-SH--E------RRRHPHPK------TA--DPA-VQIAGNF  187
AN_Q9LRR7|_AtNCED3   84 KQMNLFQRAAAAALDAAEGFLV-SH--E------K-LHPLPK------TA--DPS-VQIAGNF  127
AN_Q9C6Z1|_AtNCED5   74 SRWNPLQRAASAALDFAETALL-RR--E------R-SKPLPK------TV--DPR-HQISGNY  117
AN_O49505|_AtNCED2   67 LRLNIFQKAAAIAIDAAERALI-SH--E------Q-DSPLPK------TA--DPR-VQIAGNY  110
AN_Q9LRM7|_AtNCED6   61 PNLNLLQKLAATMLDKIESSIVIPM--E------Q-NRPLPK------PT--DPA-VQLSGNF  105
AN_O49675|_AtCCD4    73 PKLRPEMTLATALFTTVEDVIN-TF---------I-DPPSRP------SV--DPK-HVLSDNF  115
AN_O65572|_AtCCD1    16 PR--PSKGFSSKLLDLLERLVV-KL--------------MHD------AS--LPL-HYLSGNF   52
AN_Q7XJM2|_AtCCD7    45 RTGTILDESTSSA---------------------------------------FR-DYQSLFV   66
AN_Q8VY26|_AtCCD8    40 P--QVPHRLFARRIFGVTRAVI-NS--A------A-PSPLPE------KEKVEGE-RRCHVAW   83
AN_P74334|_S.ACO     17 ------------------------------------------------------PQ-DWLRGYQ   25
AN_Q9NKW9|_DmNinaB   11 ------RDFFAVK-----------------------------YDEQRN--DPQAERLDGNG   34
AN_Q6E4P5|_LeCCD1a   21 PK--PRRGITAKAIDLLEWGIV-KL--------------MHD------SS--KPL-HYLQGNF   57
An_Q8LP16|_PsNCED2   87 QKWNLLQQAAATTLDFVETTLI-KQ--E------S-KHPLPK------TS--DPR-VQIAGNF  130
AN_Q8LP15|_PsNCED3   93 SQWNPLQKAAAIALNMFESALL-SR--E------L-QYPLPK------TS--DPR-IQIAGNF  136
AN_Q9AXZ3|_PaNCED1   54 AKWNFFQRAAAVALDTVEDSFI-SGVLE------R-RHPLPK------TS--DPA-VQISGNF   99
AN_Q9AXZ4|_PaNCED3  107 PNWNIFQRTAAFALDAIEEKLI-ARVLE------R-RHPLPK------TA--DPE-VQIAGNF  152
AN_C5H805|_LeNCED2    1 ------------------------------------------------------------    0
AN_Q70YP8|_BoLCO      3 ------------------------------------------------------------    2
AN_Q84K96|_CsZCD      3 ------------------------------------------------------------    2
AN_Q8LP17|_PsCCD1    16 PK--PSNGFTSKAVDLLEKIIV-KL--------------FYD------SS--LPH-HWLSGNF   52
AN_O06785|_MtCCO      9 ------------------------------------------------SQ----N-PYLEGFL   18
AN_Q99NF1|_MmBCO2     8 ----------------------------------------------------L-PCIAPLL   15
AN_Q9JJS6|_MmBCO1     7 ------------------------------------------------------------Q-    7
AN_Q2GA76|_NaCCO      9 ----------------------------------------------------------SFTGFN   14
AN_Q53353|_PpLSD      9 --------------------------------------------------------GFSGTL   14
An_O24023|_LeNCED1   89 SKWNLVQKAAAMALDAVESALT-KH--E------L-EHPLPK------TA--DPR-VQISGNF  132
AN_Q5GN50|_GfCarX    26 ----------------------------------------------------RH-PYLTGNF   34
cons                190                                                            252

AN_Q28175|_BtRPE65   17 ------------E-TVE-ELS-SPLTAHVTGRIPLWLT-GSLLRCGPGLFEV-GS-EPFYHLF   61
AN_Q16518|_HsRPE65   17 ------------E-TVE-ELS-SPLTAHVTGRIPLWLT-GSLLRCGPGLFEV-GS-EPFYHLF   61
AN_Q9BYV7|_HsBCO2    59 ------------T-TVE-EAP-RGISARVWGHFPKWLN-GSLLRIGPGKFEF-GK-DKYNHWF  103
AN_Q9HAY6|_HsBCO1     8 --------------NRK-EQL-EPVRAKVTGKIPAWLQ-GTLLRNGPGMHTV-GE-SRYNHWF   51
AN_O24592|_ZmNCED1  128 ------------A-PVG-ERP-PVHELPVSGRIPPFID-GVYARNGANPCFD-P--VAGHHLF  171
AN_Q45VT7|_ZmCCD1    61 ------------A-PVVEETP-PAPNLTVRGHLPECLN-GEFVRVGPNPKFA-P--VAGYHWF  105
AN_Q9M9F5|_AtNCED9  188 ------------F-PVP-EKP-VVHNLPVTGTVPECIQ-GVYVRNGANPLHK-P--VSGHHLF  231
AN_Q9LRR7|_AtNCED3  128 ------------A-PVN-EQP-VRRNLPVVGKLPDSIK-GVYVRNGANPLHE-P--VTGHHFF  171
AN_Q9C6Z1|_AtNCED5  118 ------------A-PVP-EQS-VKSSLSVDGKIPDCID-GVYLRNGANPLFE-P--VSGHHLF  161
AN_O49505|_AtNCED2  111 ------------S-PVP-ESS-VRRNLTVEGTIPDCID-GVYIRNGANPMFE-P--TAGHHLF  154
AN_Q9LRM7|_AtNCED6  106 ------------A-PVN-ECP-VQNGLEVVGQIPSCLK-GVYIRNGANPMFP-P--LAGHHLF  149
AN_O49675|_AtCCD4   116 ------------A-PVLDELP-PTDCEIIHGTLPLSLN-GAYIRNGPNPQFL-P--RGPYHLF  160
AN_O65572|_AtCCD1    53 ------------A-PIRDETP-PVKDLPVHGFLPECLN-GEFVRVGPNPKFD-A--VAGYHWF   97
AN_Q7XJM2|_AtCCD7    67 ------------S-QRS-ETIEPVVIKPIEGSIPVNFPSGTYYLAGPGLFTD-DH-GSTVHPL  113
AN_Q8VY26|_AtCCD8    84 ------------T-SVQ-QEN-WEGELTVQGKIPTWLN-GTYLRNGPGLWNI-GD-HDFRHLF  128
AN_P74334|_S.ACO     26 ------------S--QP-QEW-DYWVEDVEGSIPPDLQ-GTLYRNGPGLLEI-GD-RPLKHPF   69
AN_Q9NKW9|_DmNinaB   35 RLYPNCSSDVWLRSCER-EIV-DPIEGHHSGHIPKWIC-GSLLRNGPGSWKV-GD-MTFGHLF   92
AN_Q6E4P5|_LeCCD1a   58 ------------A-PTD-ETP-PLNDLVVQGHLPECLN-GEFVRVGPNPKFA-P--VAGYHWF  101
An_Q8LP16|_PsNCED2  131 ------------A-PVP-EHP-VTQNLPITGKLPKGID-GVYLRNGANPLHE-P--VAGHHFF  174
AN_Q8LP15|_PsNCED3  137 ------------A-PVP-EQP-VVHSLPVTGKIPRCVN-GVYVRNGANPMFE-P--VSGHHLF  180
AN_Q9AXZ3|_PaNCED1  100 ------------A-PVD-EHP-VQHHLPVSGRIPRCLD-GVYLRNGANPLLE-P--VAGHHFF  143
AN_Q9AXZ4|_PaNCED3  153 ------------A-PVA-EHP-VQHGIPVAGRIPRCLD-GVYVRNGANPLFE-P--IAGHHFF  196
AN_C5H805|_LeNCED2    1 ------------------------------------------------------------F    1
AN_Q70YP8|_BoLCO      3 ------------------------------------------------------------    2
AN_Q84K96|_CsZCD      3 ------------------------------------------------------------    2
AN_Q8LP17|_PsCCD1    53 ------------A-PVKDETP-PVKDLTVQGHLPDCLN-GEFVRVGPNPKFS-P--VAGYHWF   97
AN_O06785|_MtCCO     19 ------------A-PVS-TEV-TATDLPVTGRIPEHLD-GRYLRNGPNPVAEVD--PATYHWF   63
AN_Q99NF1|_MmBCO2    16 ------------T-TAE-ETL-SAVSARVRGHIPEWLN-GYLLRVGPGKFEF-GK-DRYNHWF   60
AN_Q9JJS6|_MmBCO1     8 --------------NKK-EQL-EPVQAKVTGSIPAWLQ-GTLLRNGPGMHTV-GE-SKYNHWF   51
AN_Q2GA76|_NaCCO     15 ------------T-PSR-IEA-DIADLAHEGTIPQGLN-GAFYRVQPDPQFP-PR-LDDDIAF   59
AN_Q53353|_PpLSD     15 ------------R-PLR-IEG-DILDIEIEGEVPPQLN-GTFHRVHPDAQFP-PR-FEDDQFF   59
An_O24023|_LeNCED1  133 ------------A-PVP-ENP-VCQSLPVTGKIPKCVQ-GVYVRNGANPLFE-P--TAGHHFF  176
AN_Q5GN50|_GfCarX    35 ------------A-PIH-KTT-NLTPCTYSGCIPPELTGGQYVRNGGNPVSH-QDLGKDAHWF   81
cons                253                                                            315

AN_Q28175|_BtRPE65   62 DGQALLHKFDFK--EGHV------T-YHRRFIRTDAYVRAM--TEKRIVITE--------FGT  105
AN_Q16518|_HsRPE65   62 DGQALLHKFDFK--EGHV------T-YHRRFIRTDAYVRAM--TEKRIVITE--------FGT  105
AN_Q9BYV7|_HsBCO2   104 DGMALLHQFRMA--KGTV------T-YRSKFLQSDTYKANS--AKNRIVISE--------FGT  147
```

```
AN_Q9HAY6|_HsBCO1    52 DGLALLHSFTIR--DGEV------Y-YRSKYLRSDTYNTNI--EANRIVVSE--------FGT   95
AN_O24592|_ZmNCED1  172 DGDGMVHALRIR--NGAAE-----S-YACRFTETARLRQER--AIGRPVFPKA-------IGE  217
AN_Q45VT7|_ZmCCD1   106 DGDGMIHAMRIK--DGKA------T-YVSRYVKTARLKQEE--YFGGAKFMK--------IGD  149
AN_Q9M9F5|_AtNCED9  232 DGDGMVHAVRFD--NGSV------S-YACRFTETNRLVQER--ECGRPVFPKA-------IGE  276
AN_Q9LRR7|_AtNCED3  172 DGDGMVHAVKFE--HGSA------S-YACRFTQTNRFVQER--QLGRPVFPKA-------IGE  216
AN_Q9C6Z1|_AtNCED5  162 DGDGMVHAVKIT--NGDA------S-YSCRFTETERLVQEK--QLGSPIFPKA-------IGE  206
AN_O49505|_AtNCED2  155 DGDGMVHAVKIT--NGSA------S-YACRFTKTERLVQEK--RLGRPVFPKA-------IGE  199
AN_Q9LRM7|_AtNCED6  150 DGDGMIHAVSIGF-DNQV------S-YSCRYTKTNRLVQET--ALGRSVFPKP-------IGE  195
AN_O49675|_AtCCD4   161 DGDGMLHAIKIH--NGKA------T-LCSRYVKTYKYNVEK--QTGAPVMPNV-------FSG  205
AN_O65572|_AtCCD1    98 DGDGMIHGVRIK--DGKA------T-YVSRYVKTSRLKQEE--FFGAAKFMK--------IGD  141
AN_Q7XJM2|_AtCCD7   114 DGHGYLRAFHIDGNKRKA------T-FTAKYVKTEAKKEEH--DPVT----------------  151
AN_Q8VY26|_AtCCD8   129 DGYSTLVKLQFD--GGRI------F-AAHRLLESDAYKAAK--KHNRLCYREFSETPKSVIIN  180
AN_P74334|_S.ACO     70 DGDGMVTAFKFPG-DGRV------H-FQSKFVRTQGYVEEQ--KAGKMIYRGV-------FGS  115
AN_Q9NKW9|_DmNinaB   93 DCSALLHRFAIR--NGRV------T-YQNRFVDTETLRKNR--SAQRIVVTE--------FGT  136
AN_Q6E4P5|_LeCCD1a  102 DGDGMIHGLRIK--DGKA------T-YVSRYVRTSRLKQEE--FFGGAKFMK--------VGD  145
An_Q8LP16|_PsNCED2  175 DGDGMVHAVKFT--NGSV------S-YSCRFTETHRLAQEK--ALGRPVFPKA-------IGE  219
AN_Q8LP15|_PsNCED3  181 DGDGMVHAVTMN--DGVA------S-YACRFTETERLVQER--ELGRAMFPKA-------IGE  225
AN_Q9AXZ3|_PaNCED1  144 DGDGMVHSVSLR--QGTA------S-YACRFTETHRLVQER--AIGRPVFPKA-------IGE  188
AN_Q9AXZ4|_PaNCED3  197 DGDGMIHAVRFR--NGSA------S-YSCRYTETRRLVQER--QLSRPIFPKA-------IGE  241
AN_C5H805|_LeNCED2    2 DGDGMVHAVTVE--NGSV------S-YSCRFTETERLVQER--ELGHPVFPKA-------IGE   46
AN_Q70YP8|_BoLCO      3 ------------------------------------------------------------    2
AN_Q84K96|_CsZCD      3 ------------------------------------------------------------    2
AN_Q8LP17|_PsCCD1    98 DGDGMIHGLRIK--DGKA------S-YVSRFVKTSRFKQEE--YFNGSKFMK--------IGD  141
AN_O06785|_MtCCO     64 TGDAMVHGVALR--DGKA------RWYRNRWVRTPAVCA----ALGEPISAR----------  103
AN_Q99NF1|_MmBCO2    61 DGMALLHQFRME--RGTV------T-YKSKFLQSDTYKANS--AGGRIVISE--------FGT  104
AN_Q9JJS6|_MmBCO1    52 DGLALLHSFSIR--DGEV------F-YRSKYLQSDTYIANI--EANRIVVSE--------FGT   95
AN_Q2GA76|_NaCCO     60 NGDGMITRFHIH--DGQV------D-FRQRWAKTDKWKLEN--AAGKALFGAY-------RNP  104
AN_Q53353|_PpLSD     60 NGDGMVSLFRFH--DGKI------D-FRQRYAQTDKWKVER--KAGKSLFGAY-------RNP  104
An_O24023|_LeNCED1  177 DGDGMVHAVQFK--NGSA------S-YACRFTETERLVQEK--ALGRPVFPKA-------IGE  221
AN_Q5GN50|_GfCarX    82 DGDGMLSGVAFR--KASIDGKTIPE-FVNQYILTDLYLSRKTTSIASPIMPSI-------TTL  134
cons                316                                                              378

AN_Q28175|_BtRPE65  106 CAFPDPCK-NIFSRFFSYF--RG---------------VEVTDNALVNIYPV-GEDYYACTET  149
AN_Q16518|_HsRPE65  106 CAFPDPCK-NIFSRFFSYF--RG---------------VEVTDNALVNVYPV-GEDYYACTET  149
AN_Q9BYV7|_HsBCO2   148 LALPDPCK-NVFERFMSRFELPG-------------KAAAMTDNTNVNYVRY-KGDYYLCTET  195
AN_Q9HAY6|_HsBCO1    96 MAYPDPCK-NIFSKAFSYL--SH-------------TIPDFTDNCLINIMKC-GEDFYATSET  141
AN_O24592|_ZmNCED1  218 LHGH---S-GI-ARLALFY-ARAACG--------LVDPSAGTGVANAGLVYF-NGRLLAMSED  265
AN_Q45VT7|_ZmCCD1   150 LKGF---F-GL-FMVQMQQ-LRKKFK--------VLDFTYGFGTANTALIYH-HGKLMALSEA  197
AN_Q9M9F5|_AtNCED9  277 LHGH---L-GI-AKLMLFN-TRGLFG--------LVDPTGGLGVANAGLVYF-NGHLLAMSED  324
AN_Q9LRR7|_AtNCED3  217 LHGH---T-GI-ARLMLFY-ARAAAG--------IVDPAHGTGVANAGLVYF-NGRLLAMSED  264
AN_Q9C6Z1|_AtNCED5  207 LHGH---S-GI-ARLMLFY-ARGLFG--------LLNHKNGTGVANAGLVYF-HDRLLAMSED  254
AN_O49505|_AtNCED2  200 LHGH---S-GI-ARLMLFY-ARGLCG--------LINNQNGVGVANAGLVYF-NNRLLAMSED  247
AN_Q9LRM7|_AtNCED6  196 LHGH---S-GL-ARLALFT-ARAGIG--------LVDGTRGMGVANAGVVFF-NGRLLAMSED  243
AN_O49675|_AtCCD4   206 FNGV---TASV-ARGALTA-ARVLTG--------QYNPVNGIGLANTSLAFF-SNRLFALGES  254
AN_O65572|_AtCCD1   142 LKGF---F-GL-LMVNVQQ-LRTKLK--------ILDNTYGNGTANTALVYH-HGKLLALQEA  189
AN_Q7XJM2|_AtCCD7   152 -------------DTWRFT-HRGPFSVLK-GGKRFGNTKVMKNVANTSVLKW-AGRLLCLWEG  198
AN_Q8VY26|_AtCCD8   181 KNPF---S-GI-GEIV---------R--------LFSGESLTDNANTGVIKLGDGRVMCLTET  221
AN_P74334|_S.ACO    116 QPAG---G-W--------------LKTIF--------DLRLKNIANTNITYW-GDRLLALWEG  151
AN_Q9NKW9|_DmNinaB  137 AAVPDPCH-SIFDRFAAIF--RP---------------DSGTDNSMISIYPF-GDQYYTFTET  180
AN_Q6E4P5|_LeCCD1a  146 LKGL---F-GL-FTVYMQM-LRTKLK--------VLDISYGNSTANTALVYH-HGKLLALSEA  193
An_Q8LP16|_PsNCED2  220 LHGH---S-GI-ARLMLYY-ARSLCG--------LVDGTHGMGVANAGLVYF-NNRLLAMSED  267
AN_Q8LP15|_PsNCED3  226 LHGH---S-GI-ARLMLFY-ARSLCG--------IVDHRRGSGVANAGLVYF-DGKLLAMSED  273
AN_Q9AXZ3|_PaNCED1  189 LHGH---S-GI-ARLLLFY-ARTATG--------LVDGSSGTGVATAGLVYF-NRHLLAMSED  236
AN_Q9AXZ4|_PaNCED3  242 LHGH---S-GI-ARLLLFY-TRGLFG--------LVNADEGMGVANAGLVYF-NRRLLAMSED  289
AN_C5H805|_LeNCED2   47 LHGH---S-GI-ARLLLFY-ARGVFG--------LVDHSHGTGVANAGLVFF-NNRLLAMSED   94
AN_Q70YP8|_BoLCO      3 ---------------------------------VEPTRGIGLANTSLQFS-NGRLHALCEY   29
AN_Q84K96|_CsZCD      3 ---------------------------------VDPTKGIGLANTSLQFS-NGRLHALCEY   29
AN_Q8LP17|_PsCCD1   142 LKGL---F-GL-LMVNMQM-LRAKLK--------ILDVSYGHGTANTALVYH-HQKLLALSEG  189
AN_O06785|_MtCCO    104 PHPR---T-GI-I---------------------------EGGPNTNVLTH-AGRTLALVEA  132
AN_Q99NF1|_MmBCO2   105 LALPDPCK-SIFERFMSRF-----------------EPPTMTDNTNVNFVQY-KGDYYMSTET  148
AN_Q9JJS6|_MmBCO1    96 MAYPDPCK-NIFSKAFSYL--SH-------------TIPDFTDNCLINIMKC-GEDFYATTET  141
AN_Q2GA76|_NaCCO    105 LTDD---E-AV-K------------G--------E-----IRSTANTNAFVF-GGKLWAMKED  136
AN_Q53353|_PpLSD    105 LTDD---A-SV-Q------------G--------M-----IRGTANTNVMVH-AGKLYAMKED  136
An_O24023|_LeNCED1  222 LHGH---S-GI-ARLMLFY-ARGLFG--------LVDHSKGTGVANAGLVYF-NNRLLAMSED  269
AN_Q5GN50|_GfCarX   135 VNPL---S-TM-FQIMFAT-FRTIFLVILSNLPGSQQAIKRISVANTAVLYH-DGRALATCES  190
cons                379                                        .                *   441

AN_Q28175|_BtRPE65  150 NFITK--V-N-PETLETIKQVDLCN-----------------------YVSV----------  174
AN_Q16518|_HsRPE65  150 NFITK--I-N-PETLETIKQVDLCN-----------------------YVSV----------  174
```

```
AN_Q9BYV7|_HsBCO2   196 NFMNK--V-D-IETLEKTEKVDWSK-----------------------FIAV----------  220
AN_Q9HAY6|_HsBCO2   142 NYIRK--I-N-PQTLETLEKVDYRK-----------------------YVAV----------  166
AN_O24592|_ZmNCED1  266 DLPYHVRVAD-DGDLETVGRYDFDG-----------------------QLG-----------  292
AN_Q45VT7|_ZmCCD1   198 DKPYVVKVLE-DGDLQTLGLLDYDK-----------------------RLK-----------  224
AN_Q9M9F5|_AtNCED9  325 DLPYHVKVTQ-TGDLETSGRYDFDG-----------------------QLK-----------  351
AN_Q9LRR7|_AtNCED3  265 DLPYQVQITP-NGDLKTVGRFDFDG-----------------------QLE-----------  291
AN_Q9C6Z1|_AtNCED5  255 DLPYQVRVTD-NGDLETIGRFDFDG-----------------------QLS-----------  281
AN_O49505|_AtNCED2  248 DLPYQLKITQ-TGDLQTVGRYDFDG-----------------------QLK-----------  274
AN_Q9LRM7|_AtNCED6  244 DLPYQVKIDG-QGDLETIGRFGFDD-----------------------QID-----------  270
AN_O49675|_AtCCD4   255 DLPYAVRLTE-SGDIETIGRYDFDG-----------------------KLA-----------  281
AN_O65572|_AtCCD1   190 DKPYVIKVLE-DGDLQTLGIIDYDK-----------------------RLT-----------  216
AN_Q7XJM2|_AtCCD7   199 GEPYEI---E-SGSLDTVGRFNVENNGCESC-------------DDDDSSDRDLSGHDIWDT  243
AN_Q8VY26|_AtCCD8   222 QKGSI--LVD-HETLETIGKFEYDD-----------------------VLSD----------  247
AN_P74334|_S.ACO    152 GQPHRL---E-PSNLATIGLDDLGG-----------------------ILAEG---------  177
AN_Q9NKW9|_DmNinaB  181 PFMHR--I-N-PCTLATEARICTTD-----------------------FVGV----------  205
AN_Q6E4P5|_LeCCD1a  194 DKPYALKVLE-DGDLQTLGMLDYDK-----------------------RLT-----------  220
An_Q8LP16|_PsNCED2  268 DIPYHVRVTP-NGDLTTVCRYDFND-----------------------QLK-----------  294
AN_Q8LP15|_PsNCED3  274 DLPYELQITP-YGDLKTVGRYSFLD-----------------------QLH-----------  300
AN_Q9AXZ3|_PaNCED1  237 DLPYHVRVTS-SGDLETVGRFDFMG-----------------------QLN-----------  263
AN_Q9AXZ4|_PaNCED3  290 DLPYHVRITP-SGDLKTVGRHDFDN-----------------------QLR-----------  316
AN_C5H805|_LeNCED2   95 DVPYHVQVLP-SGDLQTVGRYNFDD-----------------------QLK-----------  121
AN_Q70YP8|_BoLCO     30 DLPYVVRLSPEDGDISTVGRIENNV-----------------------S-T-----------   56
AN_Q84K96|_CsZCD     30 DLPYVVRLSPEDGDISTVGRIENNV-----------------------S-T-----------   56
AN_Q8LP17|_PsCCD1   190 DKPYAIKVFE-DGDLQTLGMLDYDK-----------------------RLG-----------  216
AN_O06785|_MtCCO    133 GVV---NYEL-TDELDTVGPCDFDG-----------------------TLH-----------  156
AN_Q99NF1|_MmBCO2   149 NFMNK--V-D-IEMLERTEKVDWSK-----------------------FIAV----------  173
AN_Q9JJS6|_MmBCO1   142 NYIRK--I-D-PQTLETLEKVDYRK-----------------------YVAV----------  166
AN_Q2GA76|_NaCCO    137 SP-AL--VMD-PATMETFGFEKFGG-----------------------KMTG----------  161
AN_Q53353|_PpLSD    137 SP-CL--IMD-PLTLETEGYTNFDG-----------------------KLQS----------  161
An_O24023|_LeNCED1  270 DLPYHVKVTP-TGDLKTEGRFDFDG-----------------------QLK-----------  296
AN_Q5GN50|_GfCarX   191 GPPMRI---Q-LPSLDTVGWFDGVEAEGEPEISQAGSDDSPFGGSGIFSFMK-----------  238
cons                442                :                                              504

AN_Q28175|_BtRPE65  175 -----------------NGA--TAHPHIENDGTVYNIGNCFG----KNFSIAYNIV----K--  208
AN_Q16518|_HsRPE65  175 -----------------NGA--TAHPHIENDGTVYNIGNCFG----KNFSIAYNIV----K--  208
AN_Q9BYV7|_HsBCO2   221 -----------------NGA--TAHPHYDLD----------G----TAYNMGNSFG----P--  244
AN_Q9HAY6|_HsBCO1   167 -----------------NLA--TSHPHYDEA----------G----NVLNMGTSIV----E--  190
AN_O24592|_ZmNCED1  293 -----------------CAM--IAHPKLDPA---------TG----ELHALSYDV----IK--  317
AN_Q45VT7|_ZmCCD1   225 -----------------HSF--TAHPKVDPF---------TD----EMFTFGYSH----E---  248
AN_Q9M9F5|_AtNCED9  352 -----------------STM--IAHPKIDPE---------TR----ELFALSYDV----VS--  376
AN_Q9LRR7|_AtNCED3  292 -----------------STM--IAHPKVDPE---------SG----ELFALSYDV----VS--  316
AN_Q9C6Z1|_AtNCED5  282 -----------------SAM--IAHPKIDPV---------TK----ELFALSYDV----VK--  306
AN_O49505|_AtNCED2  275 -----------------SAM--IAHPKLDPV---------TK----ELHALSYDV----VK--  299
AN_Q9LRM7|_AtNCED6  271 -----------------SSV--IAHPKVDAT---------TG----DLHTLSYNV----LK--  295
AN_O49675|_AtCCD4   282 -----------------MSM--TAHPKTDPI---------TG----ETFAFRYGP----V---  305
AN_O65572|_AtCCD1   217 -----------------HSF--TAHPKVDPV---------TG----EMFTFGYSH----T---  240
AN_Q7XJM2|_AtCCD7   244 AADLLKPILQGVFKMPPKRF--LSHYKVDGR---------RK----RLLTVTCNAEDMLLP--  289
AN_Q8VY26|_AtCCD8   248 ------------------HMIQSAHPIVTE-----------T----EMWTLIPDL----VK--  271
AN_P74334|_S.ACO    178 -----------------QPL--SAHPRIDPA---------STFDGGQPCYVTFSIKS--SL--  208
AN_Q9NKW9|_DmNinaB  206 -----------------VNH--TSHPHVLPS----------G----TVYNLGTTMT----R--  229
AN_Q6E4P5|_LeCCD1a  221 -----------------HSF--TAHPKVDAV---------TG----EMFTFGYAH----T---  244
An_Q8LP16|_PsNCED2  295 -----------------STM--IAHPKVDPV---------DK----NLYALSYDV----VQ--  319
AN_Q8LP15|_PsNCED3  301 -----------------STM--IAHPKIDPI---------SG----ELFALSYEV----A---  324
AN_Q9AXZ3|_PaNCED1  264 -----------------SAM--IAHPKLDPA---------SG----ELFALSYNV----IK--  288
AN_Q9AXZ4|_PaNCED3  317 -----------------SSM--IAHPKLDPE---------SG----ELFSLSYDV----AR--  341
AN_C5H805|_LeNCED2  122 -----------------STM--IAHPKIDPV---------SG----ELFALSYDV----VQ--  146
AN_Q70YP8|_BoLCO     57 -----------------KST--TAHPKTDPV---------TG----ETFSFSYGP----I---   80
AN_Q84K96|_CsZCD     57 -----------------KST--TAHPKTDPV---------TG----ETFSFSYGP----I---   80
AN_Q8LP17|_PsCCD1   217 -----------------HNF--TAHPKVDPF---------TG----EMFTFGYSH----T---  240
AN_O06785|_MtCCO    157 -----------------GGY--TAHPQRDPH---------TG----ELHAVSYSF----AR--  181
AN_Q99NF1|_MmBCO2   174 -----------------NGA--TAHPHYDPD----------G----TAYNMGNSYG----P--  197
AN_Q9JJS6|_MmBCO1   167 -----------------NLA--TSHPHYDEA----------G----NVLNMGTSVV----D--  190
AN_Q2GA76|_NaCCO    162 -----------------QTF--TAHPKVDPK---------TG----NMVAIGYAA----SGLC  188
AN_Q53353|_PpLSD    162 -----------------QTF--CAHPKIDPV---------TG----NLCAFAYGA----KGLM  188
An_O24023|_LeNCED1  297 -----------------STM--IAHPKLDPV---------SG----ELFALSYDV----IQ--  321
AN_Q5GN50|_GfCarX   239 -----------------EWT--TGHPKVDPV---------TG----EMLLYHNT-----FM--  262
cons                505                      .*                                     567

AN_Q28175|_BtRPE65  209 IPPLQADKED-----PISKS----------EIV------VQ--FPCSDRFKPSYVHSFGLTPN  248
```

```
AN_Q16518|_HsRPE65  209 IPPLQADKED-----PISKS----------EIV------VQ--FPCSDRFKPSYVHSFGLTPN  248
AN_Q9BYV7|_HsBCO2   245 YG-FSYKVIRVP---PEKVDL-GE-TIHGVQVI------CS--IASTEKGKPSYYHSFGMTRN  293
AN_Q9HAY6|_HsBCO1   191 KGKTKYVIFKIPATVPEGKKQ-GKSPWKHTEVF------CS--IPSRSLLSPSYYHSFGVTEN  244
AN_O24592|_ZmNCED1  318 RPYLKYFYFR-----PDGTKS-D---------D------VE--I---PLEQPTMIHDFAITEN  354
AN_Q45VT7|_ZmCCD1   249 PPYCTYRVIN-----KEGAML-D---------P------VP--I---TIPESVMMHDFAITEN  285
AN_Q9M9F5|_AtNCED9  377 KPYLKYFRFT-----SDGEKS-P---------D------VE--I---PLDQPTMIHDFAITEN  413
AN_Q9LRR7|_AtNCED3  317 KPYLKYFRFS-----PDGTKS-P---------D------VE--I---QLDQPTMMHDFAITEN  353
AN_Q9C6Z1|_AtNCED5  307 KPYLKYFKFS-----PEGEKS-P---------D------VE--I---PLASPTMMHDFAITEN  343
AN_O49505|_AtNCED2  300 KPYLKYFRFS-----PDGVKS-P---------E------LE--I---PLETPTMIHDFAITEN  336
AN_Q9LRM7|_AtNCED6  296 KPHLRYLKFN-----TCGKKT-R---------D------VE--I---TLPEPTMIHDFAITEN  332
AN_O49675|_AtCCD4   306 PPFLTYFRFD-----SAGKKQ-R---------D------VP--I--FSMTSPSFLHDFAITKR  343
AN_O65572|_AtCCD1   241 PPYLTYRVIS-----KDGIMH-D---------P------VP--I---TISEPIMMHDFAITET  277
AN_Q7XJM2|_AtCCD7   290 RSNFTFCEYD-----SEFKLI-Q---------T------KE--F---KIDDHMMIHDWAFTDT  326
AN_Q8VY26|_AtCCD8   272 -PGYRVVRME-----AGSNKR-E------VVGR------VRCRS---GSWGPGWVHSFAVTEN  312
AN_P74334|_S.ACO    209 SSTLTLLELD-----PQGKLL-R---------Q------KT--E---TFPGFAFIHDFAITPH  245
AN_Q9NKW9|_DmNinaB  230 SG-PAYTILCFP---HGEQMF-ED-----AHVV------AT--LPCRWKLHPGYMHTFGLTDH  274
AN_Q6E4P5|_LeCCD1a  245 PPYITYRVIS-----KDGIMQ-D---------P------VP--I---TIPEPIMMHDFAITEN  281
An_Q8LP16|_PsNCED2  320 KPYLKYFRFD-----SNGVKS-P---------D------VE--I---PLAEPTMMHDFAITEN  356
AN_Q8LP15|_PsNCED3  325 RPYLKYFRFS-----PDGKKS-P---------D------VE--I---RLSVPTMIHDFAITEN  361
AN_Q9AXZ3|_PaNCED1  289 KPFLKFFKFT-----SDGKKS-P---------D------VE--I---PIDQPTMIHDFVITEN  325
AN_Q9AXZ4|_PaNCED3  342 KPYLKYFHFA-----PDGWKS-P---------D------VE--I---PLDRPTMIHDFAITEN  378
AN_C5H805|_LeNCED2  147 KPYLKSFKFS-----PDGEKS-P---------D------VE--I---PLDVPTMMHDFAITES  183
AN_Q70YP8|_BoLCO     81 QPYVTYSRYD-----CDDKKSGP---------D------VP--I--FSFKEPSFVHDFAITEH  119
AN_Q84K96|_CsZCD     81 QPYVTYSRYD-----CDGKKSGP---------D------VP--I--FSFKEPSFVHDFAITEH  119
AN_Q8LP17|_PsCCD1   241 APYVTYRVIS-----KDGFMQ-D---------P------VP--I---TISDPVMMHDFAITEN  277
AN_O06785|_MtCCO    182 GHRVQYSVIG-----TDGHAR-R---------T------VD--I---EVAGSPMMHSFSLTDN  218
AN_Q99NF1|_MmBCO2   198 RG-SCYNIIRVP---PKKKEP-GE-TIHGAQVL------CS--IASTEKMKPSYYHSFGMTKN  246
AN_Q9JJS6|_MmBCO1   191 KGRTKYVIFKIPATVPDSKKK-GKSPVKHAEVF------CS--ISSRSLLSPSYYHSFGVTEN  244
AN_Q2GA76|_NaCCO    189 TDDVTYMEVS-----PEGELV-R---------E------VW--F---KVPYYCMMHDFGITED  225
AN_Q53353|_PpLSD    189 TLDMAYIEIS-----PTGKLL-K---------E------IP--F---QNPYYCMMHDFGVTED  225
An_O24023|_LeNCED1  322 KPYLKYFRFS-----KNGEKS-N---------D------VE--I---PVEDPTMMHDFAITEN  358
AN_Q5GN50|_GfCarX   263 PPYVHCSVLP-----KSNEKA-P---------GHRLVNQPV--L---GVSGARMMHDFGASRS  305
cons                568                                                  * :  :    630

AN_Q28175|_BtRPE65  249 YIVFVETPVKINLF--K-FL---SS-WSLWGANYMD--------------------------  277
AN_Q16518|_HsRPE65  249 YIVFVETPVKINLF--K-FL---SS-WSLWGANYMD--------------------------  277
AN_Q9BYV7|_HsBCO2   294 YIIFIEQPLKMNLW--K-IA---TS-KI-RGKAFSD--------------------------  321
AN_Q9HAY6|_HsBCO1   245 YVIFLEQPFRLDIL--K-MA---TA-YI-RRMSWAS--------------------------  272
AN_O24592|_ZmNCED1  355 LVVVPDHQVVFKLQ--E-ML---RG-G--------S--------------------------  375
AN_Q45VT7|_ZmCCD1   286 YSIFMDLPLLFRPK--E-MV---KN-G--------E--------------------------  306
AN_Q9M9F5|_AtNCED9  414 FVVIPDQQVVFRLP--E-MI---RG-G--------S--------------------------  434
AN_Q9LRR7|_AtNCED3  354 FVVVPDQQVVFKLP--E-MI---RG-G--------S--------------------------  374
AN_Q9C6Z1|_AtNCED5  344 FVVIPDQQVVFKLS--D-MF---LG-K--------S--------------------------  364
AN_O49505|_AtNCED2  337 FVVIPDQQVVFKLG--E-MI---SG-K--------S--------------------------  357
AN_Q9LRM7|_AtNCED6  333 FVVIPDQQMVFKLS--E-MI---RG-G--------S--------------------------  353
AN_O49675|_AtCCD4   344 HAIFAEIQLGMRMNMLDLVL---EG-G--------S--------------------------  367
AN_O65572|_AtCCD1   278 YAIFMDLPMHFRPK--E-MV---KE-K--------K--------------------------  298
AN_Q7XJM2|_AtCCD7   327 HYILFANRVKLNPI--G-SIAAMCG-M--------S--------------------------  350
AN_Q8VY26|_AtCCD8   313 YVVIPEMPLRYSVK--N-LL---RA-E--------P--------------------------  333
AN_P74334|_S.ACO    246 YAIFLQNNVTLNGL--P-YLFGLRGAG--------E--------------------------  270
AN_Q9NKW9|_DmNinaB  275 YFVIVEQPLSVSLT--E-YI---KA-QL-GGQNLSA--------------------------  302
AN_Q6E4P5|_LeCCD1a  282 YAIMMDLPLCFRPK--E-MV---KK-N--------K--------------------------  302
An_Q8LP16|_PsNCED2  357 FVVVPDQQVVFKLG--E-MI---RG-G--------S--------------------------  377
AN_Q8LP15|_PsNCED3  362 YVVIPDHQVVFKLE--E-MI---KG-C--------S--------------------------  382
AN_Q9AXZ3|_PaNCED1  326 FVIIPDQQVVFKLQ--E-MI---RG-G--------S--------------------------  346
AN_Q9AXZ4|_PaNCED3  379 FVVIPDQQVVFKLE--E-MI---RG-G--------S--------------------------  399
AN_C5H805|_LeNCED2  184 YVVIPDQQVVFKLQ--E-MI---KG-G--------S--------------------------  204
AN_Q70YP8|_BoLCO    120 YAVFPDIQIVMKPA--E-IV---RG-R--------R--------------------------  140
AN_Q84K96|_CsZCD    120 YAVFPDIQIVMKPA--E-IV---RG-R--------R--------------------------  140
AN_Q8LP17|_PsCCD1   278 YSIFMDLPLYFRPK--E-MV---KN-K--------T--------------------------  298
AN_O06785|_MtCCO    219 YVVIYDLPVTFDPM--Q-VV---PA-S--------VPRWLQRPARLVIQSVLGRVRIPDPIAA  266
AN_Q99NF1|_MmBCO2   247 YIIFVEQPVKMKLW--K-II---TS-KI-RGKPFAD--------------------------  274
AN_Q9JJS6|_MmBCO1   245 YVVFLEQPFKLDIL--K-MA---TA-YM-RGVSWAS--------------------------  272
AN_Q2GA76|_NaCCO    226 YLVLHIVPSIGSWE--R-LE---QG-K--------P--------------------------  246
AN_Q53353|_PpLSD    226 YAVFAVMPLLSSWD--R-LE---QR-L--------P--------------------------  246
An_O24023|_LeNCED1  359 FVVIPDQQVVFKMS--E-MI---RG-G--------S--------------------------  379
AN_Q5GN50|_GfCarX   306 HTIIMDLPLSLDPL--N-TM---KG-K--------E--------------------------  326
cons                631   :.                                                       693
```

```
AN_Q28175|_BtRPE65  278 ------------C----FESNE----TMGVWLHIADKKR-KKYINN---KY----R--TSPFN  310
AN_Q16518|_HsRPE65  278 ------------C----FESNE----TMGVWLHIADKKR-KKYLNN---KY----R--TSPFN  310
AN_Q9BYV7|_HsBCO2   322 ------------G----ISWEP----QCNTRFHVVEKRT-GQLLPG---RY----Y--SKPFV  354
AN_Q9HAY6|_HsBCO1   273 ------------C----LAFHR----EEKTYIHIIDQRT-RQPVQT---KF----Y--TDAMV  305
AN_O24592|_ZmNCED1  376 ------------P----VVLD----KEKTSRFGVLPKHA-ADASEM---AW----VD-VPDCF  409
AN_Q45VT7|_ZmCCD1   307 ------------FI---YKFD----PTKKGRFGILPRYA-KDDKLI---RW----FQ-LPNCF  341
AN_Q9M9F5|_AtNCED9  435 ------------P----VVYD----EKKKSRFGILNKNA-KDASSI---QW----IE-VPDCF  468
AN_Q9LRR7|_AtNCED3  375 ------------P----VVYD----KNKVARFGILDKYA-EDSSNI---KW----ID-APDCF  408
AN_Q9C6Z1|_AtNCED5  365 ------------P----VKYD----GEKISRFGILPRNA-KDASEM---VW----VE-SPETF  398
AN_O49505|_AtNCED2  358 ------------P----VVFD----GEKVSRLGIMPKDA-TEASQI---IW----VN-SPETF  391
AN_Q9LRM7|_AtNCED6  354 ------------P----VIYV----KEKMARFGVLSKQD-LTGSDI---NW----VD-VPDCF  387
AN_O49675|_AtCCD4   368 ------------P----VGTD----NGKTPRLGVIPKYA-GDESEM---KW----FE-VPGFN  401
AN_O65572|_AtCCD1   299 ------------MI---YSFD----PTKKARFGVLPRYA-KDELMI---RW----FE-LPNCF  333
AN_Q7XJM2|_AtCCD7   351 ------------PMVSALSLNP---SNESSPIYILPRFS-D--KYSRGGRDWRVPVEVSSQLW  395
AN_Q8VY26|_AtCCD8   334 ------------T----PLYKFEWCPQDGAFIHVMSKL-T--GEVV---AS----VE-VPAYV  369
AN_P74334|_S.ACO    271 ------------C----VQFHP----DKPAQIILVPRDG----GEI---KR----IP-VQAGF  301
AN_Q9NKW9|_DmNinaB  303 ------------C----LKWFE----DRPTLFHLIDRVS-GK--LV---QT----YE-SEAFF  334
AN_Q6E4P5|_LeCCD1   303 ------------LA---FTFD----ATKKARFGVLPRYA-NNEALI---RW----FE-LPNCF  337
An_Q8LP16|_PsNCED2  378 ------------P----VVYD----KEKVSRFGVLSKNA-ENASEM---KW----ID-APECF  411
AN_Q8LP15|_PsNCED3  383 ------------P----VIFD----GAKKSRFGVLPKYA-KDASSI---IW----VD-SPDTF  416
AN_Q9AXZ3|_PaNCED1  347 ------------P----VVYD----KKKTARFGILLKTA-ADSNGL---RW----ID-APDCF  380
AN_Q9AXZ4|_PaNCED3  400 ------------P----VVYD----KNKTSRFGILPKYA-PDASEM---IW----VD-APDCF  433
AN_C5H805|_LeNCED2  205 ------------P----VIYD----KNKKSRFGILPKNA-ENSENI---IW----VE-SAETF  238
AN_Q70YP8|_BoLCO    141 ------------M----IGPD----LEKVPRLGLLPRYA-TSDSEM---RW----FD-VPGFN  174
AN_Q84K96|_CsZCD    141 ------------M----IGPD----LEKVPRLGLLPRYA-TSDSEM---RW----FD-VPGFN  174
AN_Q8LP17|_PsCCD1   299 ------------LI---FSFD----STKKARFGVLPRYA-KDDKHI---RW----FE-LPNCF  333
AN_O06785|_MtCCO    267 LGNRMQGHSDRLP----YAWN----PSYPARVGVMPREG-GN-EDV---RW----FD-IEPCY  311
AN_Q99NF1|_MmBCO2   275 ------------G----ISWEP----QYNTRFHVVDKHT-GQLLPG---MY----Y--SMPFL  307
AN_Q9JJS6|_MmBCO1   273 ------------C----MSFDR----EDKTYIHIIDQRT-RKPVPT---KF----Y--TDPMV  305
AN_Q2GA76|_NaCCO    247 ------------H----FGFD----TTMPVHLGIIPRRDGVRQEDI---RW----FT-RDNCF  281
AN_Q53353|_PpLSD    247 ------------F----FGFD----TTLPCYLGILPRN-G-DARDL---RW----FK-TGNCF  279
An_O24023|_LeNCED1  380 ------------P----VVYD----KNKVSRFGILDKYA-KDGSDL---KW----VE-VPDCF  413
AN_Q5GN50|_GfCarX   327 ------------V----VAYDP----TKPSRFGVFPRHL-P--SSV---RW----FH-TAPCC  358
cons                694                                 . :  :                          756

AN_Q28175|_BtRPE65  311 LFHHINTYEDHE--------F-L-IV--D-LCC-------------------WKGFEF-V---  337
AN_Q16518|_HsRPE65  311 LFHHINTYEDNG--------F-L-IV--D-LCC-------------------WKGFEF-V---  337
AN_Q9BYV7|_HsBCO2   355 TFHQINAFEDQG--------C-V-II--D-LCC-------------------QDNGRT-L---  381
AN_Q9HAY6|_HsBCO1   306 VFHHVNAYEEDG--------C-I-VF--D-VIA-------------------YEDNSL-Y---  332
AN_O24592|_ZmNCED1  410 CFHLWNAWEDEA--T-G---E-V-VV--I-GSC-------------------MTPADS-I---  438
AN_Q45VT7|_ZmCCD1   342 IFHNANAWEEGD--------E-V-VL--I-TCR-------------------LENPDLDKVNG  372
AN_Q9M9F5|_AtNCED9  469 CFHLWNSWEEPE--T-D---E-V-VV--I-GSC-------------------MTPPDS-I---  497
AN_Q9LRR7|_AtNCED3  409 CFHLWNAWEEPE--T-D---E-V-VV--I-GSC-------------------MTPPDS-I---  437
AN_Q9C6Z1|_AtNCED5  399 CFHLWNAWESPE--T-D---E-V-VV--I-GSC-------------------MTPADS-I---  427
AN_O49505|_AtNCED2  392 CFHLWNAWESPE--T-E---E-I-VV--I-GSC-------------------MSPADS-I---  420
AN_Q9LRM7|_AtNCED6  388 CFHLWNAWEERTEEG-D---PVI-VV--I-GSC-------------------MSPPDT-I---  419
AN_O49675|_AtCCD4   402 IIHAINAWDEDD--G-N---S-V-VL--I-APN-------------------IMSIEH-T---  430
AN_O65572|_AtCCD1   334 IFHNANAWEEED--------E-V-VL--I-TCR-------------------LENPDLDMVSG  364
AN_Q7XJM2|_AtCCD7   396 LIHSGNAYETRE--DNG---D-L-KIQIQASACSYRWFDFQKMFGYDWQSNKLDPSVM-N---  447
AN_Q8VY26|_AtCCD8   370 TFHFINAYEEDK--N-GDGKA-TVII--A-DCC-------------------EHNADTRI---  403
AN_P74334|_S.ACO    302 VFHHANAFEEN-----G---K-I-IL--D-SIC-------------------YNSLPQ-V---  328
AN_Q9NKW9|_DmNinaB  335 YLHIINCFERDG--------H-V-VV--D-ICS-------------------YRNPEM-I---  361
AN_Q6E4P5|_LeCCD1a  338 IFHNANAWEEGD--------E-V-VL--I-TCR-------------------LVNPDLDMVNG  368
An_Q8LP16|_PsNCED2  412 CFHLWNAWEEPE--N-D---E-V-VV--I-GSC-------------------MTPADS-I---  440
AN_Q8LP15|_PsNCED3  417 CFHLWNAWEEPE--A-D---E-I-VV--I-GSC-------------------MTPPDS-I---  445
AN_Q9AXZ3|_PaNCED1  381 CFHLWTAWEEPE--T-D---Q-V-VV--I-GSC-------------------MTPPDS-I---  409
AN_Q9AXZ4|_PaNCED3  434 CFHLWNAWEEPE--S-G---E-V-VV--V-GSC-------------------MTPPDS-I---  462
AN_C5H805|_LeNCED2  239 CFHLWNAWEEPE--T-D---E-V-IV--I-GSC-------------------MTPPDS-I---  267
AN_Q70YP8|_BoLCO    175 MVHVVNAWEEEG--G-E---V-V-VI--V-APN-------------------VSPIEN-A---  203
AN_Q84K96|_CsZCD    175 MVHVVNAWEEEG--G-E---V-V-VI--V-APN-------------------VSPIEN-A---  203
AN_Q8LP17|_PsCCD1   334 IFHNANAWEEED--------E-I-VL--I-TCR-------------------LENPNLDVVGG  364
AN_O06785|_MtCCO    312 VYHPLNAYSECR--N-G--AE-V-LV--L-DVV-------------------RYS--R-M---  339
AN_Q99NF1|_MmBCO2   308 TYHQINAFEDQG--------C-I-VI--D-LCC-------------------QDDGRS-L---  334
AN_Q9JJS6|_MmBCO1   306 VFHHVNAYEEDG--------C-V-LF--D-VIA-------------------YEDSSL-Y---  332
AN_Q2GA76|_NaCCO    282 ASHVLNAWQEGT--K--------IHF--V-TCE-------------------AKNNMFPF---  309
AN_Q53353|_PpLSD    280 VGHVMNAFNDGT--K--------VHI--D-MPV-------------------SRNNSFPF---  307
An_O24023|_LeNCED1  414 CFHLWNAWEEAE--T-D---E-I-VV--I-GSC-------------------MTPPDS-I---  442
AN_Q5GN50|_GfCarX   359 IFHTANTWDSQS--SEG---E-L-SVNLL-ACR-------------------MTSSTL-V---  390
cons                757   *  . :.              .                                      819
```

```
AN_Q28175|_BtRPE65  338 -Y-NYS---------YLA-------------NLR----------------------------  347
AN_Q16518|_HsRPE65  338 -Y-NYL---------YLA-------------NLR----------------------------  347
AN_Q9BYV7|_HsBCO2   382 ---EVY---------QLQ-------------NLR----------------------------  390
AN_Q9HAY6|_HsBCO1   333 ---QLF---------YLA-------------NLN----------------------------  341
AN_O24592|_ZmNCED1  439 -F-N------E----SDE-------------RLE----------------------------  447
AN_Q45VT7|_ZmCCD1   373 YQ-S------D----KLE-------------NFG----------------------------  382
AN_Q9M9F5|_AtNCED9  498 -F-N------E----HDE-------------TLQ----------------------------  506
AN_Q9LRR7|_AtNCED3  438 -F-N------E----SDE-------------NLK----------------------------  446
AN_Q9C6Z1|_AtNCED5  428 -F-N------E----CDE-------------QLN----------------------------  436
AN_O49505|_AtNCED2  421 -F-N------E----RDE-------------SLR----------------------------  429
AN_Q9LRM7|_AtNCED6  420 -F-S------E----SGE-------------PTR----------------------------  428
AN_O49675|_AtCCD4   431 -L-E-----------RMD-------------LVH----------------------------  438
AN_O65572|_AtCCD1   365 KV-K------E----KLE-------------NFG----------------------------  374
AN_Q7XJM2|_AtCCD7   448 -L-NR----------GDD-------------KLL----------------------------  456
AN_Q8VY26|_AtCCD8   404 -L-DML---------RLD-------------TLR----------------------------  413
AN_P74334|_S.ACO    329 -D-TD----------GDFR----ST---NFDNLD----------------------------  343
AN_Q9NKW9|_DmNinaB  362 ---NCM---------YLE-------------AIA----------------------------  370
AN_Q6E4P5|_LeCCD1a  369 AV-K------E----KLE-------------NFC----------------------------  378
An_Q8LP16|_PsNCED2  441 -F-N------E----CDE-------------SLK----------------------------  449
AN_Q8LP15|_PsNCED3  446 -F-N------E----SDE-------------NLK----------------------------  454
AN_Q9AXZ3|_PaNCED1  410 -F-N------E----SNQ-------------SLK----------------------------  418
AN_Q9AXZ4|_PaNCED3  463 -F-N------E----NEE-------------SLK----------------------------  471
AN_C5H805|_LeNCED2  268 -F-N------E----CNE-------------NLK----------------------------  276
AN_Q70YP8|_BoLCO    204 -I-D-----------RFD-------------LLH----------------------------  211
AN_Q84K96|_CsZCD    204 -I-D-----------RFD-------------LLH----------------------------  211
AN_Q8LP17|_PsCCD1   365 AV-K------E----KLD-------------NFS----------------------------  374
AN_O06785|_MtCCO    340 -F-D------RDRRGPGG-------------DSR----------------------------  352
AN_Q99NF1|_MmBCO2   335 ---DLY---------QLQ-------------NLR----------------------------  343
AN_Q9JJS6|_MmBCO1   333 ---QLF---------YLA-------------NLN----------------------------  341
AN_Q2GA76|_NaCCO    310 -FPDVHGAPFN----GME-------------AM-----------------------------  324
AN_Q53353|_PpLSD    308 -F-DVHGA-------PFD-------------PVA----------------------------  319
An_O24023|_LeNCED1  443 -F-N------E----CDE-------------GLK----------------------------  451
AN_Q5GN50|_GfCarX   391 -Y-TA----------GNIRPPVRSRCTQARVWSDEREETACRYKEAPALESPGESTGLADYFP  441
cons                820                                                               882

AN_Q28175|_BtRPE65  348 ----------------------EN---WEEVKKNARK--A-PQPEVR------RYVLPLNI--  374
AN_Q16518|_HsRPE65  348 ----------------------EN---WEEVKKNARK--A-PQPEVR------RYVLPLNI--  374
AN_Q9BYV7|_HsBCO2   391 ----------------------KAGEGLDQV-HNSAA--K---SFPR------RFVLPLNV--  417
AN_Q9HAY6|_HsBCO1   342 ----------------------QD---FKEN-SRLTS--V---PTLR------RFAVPLHV--  365
AN_O24592|_ZmNCED1  448 ------------S----------V---LTEIRLDART--G-R--STR------R---------  465
AN_Q45VT7|_ZmCCD1   383 ------------N----------E---LYEMRFNMKT--G-A--ASQ------K---------  400
AN_Q9M9F5|_AtNCED9  507 ------------S----------V---LSEIRLNLKT--G-E--STR------R---------  524
AN_Q9LRR7|_AtNCED3  447 ------------S----------V---LSEIRLNLKT--G-E--STR------R---------  464
AN_Q9C6Z1|_AtNCED5  437 ------------S----------V---LSEIRLNLKT--G-K--STR------R---------  454
AN_O49505|_AtNCED2  430 ------------S----------V---LSEIRINLRT--R-K--TTR------R---------  447
AN_Q9LRM7|_AtNCED6  429 ------------V----------E---LSEIRLNMRT--K-E--SNR------K---------  446
AN_O49675|_AtCCD4   439 ------------A----------L---VEKVKIDLVT--G-I--VRR------H---------  456
AN_O65572|_AtCCD1   375 ------------N----------E---LYEMRFNMKT--G-S--ASQ------K---------  392
AN_Q7XJM2|_AtCCD7   457 ------------P----------H---LVKVSMTLDS-TG-N--CNS------C---------  475
AN_Q8VY26|_AtCCD8   414 ------------SSHGHDVLPDAR---IGRFRIPLDG--S-K--YGKLETAVEA---------  447
AN_P74334|_S.ACO    344 ------------PG----------Q---LWRFTI--DPAAA-T--VEK------Q---------  362
AN_Q9NKW9|_DmNinaB  371 ----------------------NMQ--TNPN-YATLF--R---GRPL------RFVLPLGTIP  397
AN_Q6E4P5|_LeCCD1a  379 ------------N----------E---LYEMRFNMKS--G-A--ASQ------K---------  396
An_Q8LP16|_PsNCED2  450 ------------S----------V---LSEIRLNLKT--G-K--STR------R---------  467
AN_Q8LP15|_PsNCED3  455 ------------S----------V---LSEIRLNLKT--G-K--STR------R---------  472
AN_Q9AXZ3|_PaNCED1  419 ------------S----------V---LTEIRLNLKT--G-L--SSR------R---------  436
AN_Q9AXZ4|_PaNCED3  472 ------------S----------I---LTEIRLNTRT--G-E--STR------R---------  489
AN_C5H805|_LeNCED2  277 ------------S----------V---LSEIRLNLKT--G-E--STR------R---------  294
AN_Q70YP8|_BoLCO    212 ------------V----------S---VEMARIELKS--G-S--VPR------T---------  229
AN_Q84K96|_CsZCD    212 ------------V----------S---VEMARIELKS--G-S--VSR------T---------  229
AN_Q8LP17|_PsCCD1   375 ------------N----------E---LYEMRFNMKT--G-E--ASQ------K---------  392
AN_O06785|_MtCCO    353 ------------P----------S---LDRWTINLAT--G-A--VTA------E---------  370
AN_Q99NF1|_MmBCO2   344 ----------------------KAGEGLDQV-YELKA--K---SFPR------RFVLPLDV--  370
AN_Q9JJS6|_MmBCO1   342 ----------------------KD---FEEK-SRLTS--V---PTLR------RFAVPLHV--  365
AN_Q2GA76|_NaCCO    325 ----------------------SH---PTDWVVDMAS--NGE--DFA-G---IV---------  345
AN_Q53353|_PpLSD    320 ------------G--------QGF---LTRWTVDMAS--NGD--SFE-K---TE---------  342
An_O24023|_LeNCED1  452 ------------S----------V---LSEIRLNLKT--G-K--STR------K---------  469
AN_Q5GN50|_GfCarX   442 ITAESDDYDQCRL----------Y---YYEFDLAMESRNH-V--KSQ------W---------  473
```

```
cons                883                                                                  945

AN_Q28175|_BtRPE65  375 ----------------DK----ADTGK-----------------------------------  381
AN_Q16518|_HsRPE65  375 ----------------DK----ADTGK-----------------------------------  381
AN_Q9BYV7|_HsBCO2   418 ----------------SLN---APEGD-----------------------------------  425
AN_Q9HAY6|_HsBCO1   366 ----------------DKN---AEVGT-----------------------------------  373
AN_O24592|_ZmNCED1  466 -------AVLP-P-SQ-------------------------------------------------  472
AN_Q45VT7|_ZmCCD1   401 -------QLS-------------------------------------------------------  403
AN_Q9M9F5|_AtNCED9  525 -------PVIS----E-------------------------------------------------  529
AN_Q9LRR7|_AtNCED3  465 -------PIISNE-DQ-------------------------------------------------  472
AN_Q9C6Z1|_AtNCED5  455 -------TIIP-G-SV-------------------------------------------------  461
AN_O49505|_AtNCED2  448 -------SLLV-N--E-------------------------------------------------  453
AN_Q9LRM7|_AtNCED6  447 -------VIVT-----------------------------------------------------  450
AN_O49675|_AtCCD4   457 -------PIS-------------------------------------------------------  459
AN_O65572|_AtCCD1   393 -------KLS-------------------------------------------------------  395
AN_Q7XJM2|_AtCCD7   476 -------DVEP-LNGW-------------------------------------------------  483
AN_Q8VY26|_AtCCD8   448 -------EKH-------------------------------------------------------  450
AN_P74334|_S.ACO    363 -------------LMV-------------------------------------------------  365
AN_Q9NKW9|_DmNinaB  398 PASIAKRGLVK-S-FSLAGLSAPQVSRTMKHSVSQYADITYMPTNGKQATAGEESPKRDAKRG  458
AN_Q6E4P5|_LeCCD1   397 -------KLS-------------------------------------------------------  399
An_Q8LP16|_PsNCED2  468 -------AIIQ-E-SE-------------------------------------------------  474
AN_Q8LP15|_PsNCED3  473 -------SIVP-----------------------------------------------------  476
AN_Q9AXZ3|_PaNCED1  437 -------EIDP---SR-------------------------------------------------  442
AN_Q9AXZ4|_PaNCED3  490 -------TIID-P-QK-------------------------------------------------  496
AN_C5H805|_LeNCED2  295 -------QLPS-P-SD-------------------------------------------------  301
AN_Q70YP8|_BoLCO    230 -------LLS-------------------------------------------------------  232
AN_Q84K96|_CsZCD    230 -------LLS-------------------------------------------------------  232
AN_Q8LP17|_PsCCD1   393 -------KLS-------------------------------------------------------  395
AN_O06785|_MtCCO    371 -------CRD-------------------------------------------------------  373
AN_Q99NF1|_MmBCO2   371 ----------------SVD---AAEGK-----------------------------------  378
AN_Q9JJS6|_MmBCO1   366 ----------------DKD---AEVGS-----------------------------------  373
AN_Q2GA76|_NaCCO    346 -------KLS-------------------------------------------------------  348
AN_Q53353|_PpLSD    343 -------RLF-------------------------------------------------------  345
An_O24023|_LeNCED1  470 -------SIIENP-DE-------------------------------------------------  477
AN_Q5GN50|_GfCarX   474 -------------ALS-------------------------------------------------  476
cons                946                                                                 1008

AN_Q28175|_BtRPE65  382 --------NLVTLPNTTATAILCSDETIWLEPEVLFS---GPRQAFEFPQINYQKYGGKPYTY  433
AN_Q16518|_HsRPE65  382 --------NLVTLPNTTATAILCSDETIWLEPEVLFS---GPRQAFEFPQINYQKYCGKPYTY  433
AN_Q9BYV7|_HsBCO2   426 --------NLSPLSYTSASAVKQADGTIWCSHENLHQEDLEKEGGIEFPQIYYDRFSGKKYHF  480
AN_Q9HAY6|_HsBCO1   374 --------NLIKVASTTATALKEEDGQVYCQP-------EFLYEGLELPRVNYA-HNGKQYRY  420
AN_O24592|_ZmNCED1  473 ----------------------------------------QVNLEVGMVNRN-LLGRETRY  492
AN_Q45VT7|_ZmCCD1   404 ----------------------------------------VSAVDFPRVNES-YTGRKQRY  423
AN_Q9M9F5|_AtNCED9  530 ----------------------------------------QVNLEAGMVNRN-LLGRKTRY  549
AN_Q9LRR7|_AtNCED3  473 ----------------------------------------QVNLEAGMVNRN-MLGRKTKF  492
AN_Q9C6Z1|_AtNCED5  462 ----------------------------------------QMNLEAGMVNRN-LLGRKTRY  481
AN_O49505|_AtNCED2  454 ----------------------------------------DVNLEIGMVNRN-RLGRKTRF  473
AN_Q9LRM7|_AtNCED6  451 ----------------------------------------GVNLEAGHINRS-YVGRKSQF  470
AN_O49675|_AtCCD4   460 ----------------------------------------ARNLDFAVINPA-FLGRCSRY  479
AN_O65572|_AtCCD1   396 ----------------------------------------ASAVDFPRINEC-YTGKKQRY  415
AN_Q7XJM2|_AtCCD7   484 ----------------------------------------NKPSDFPVINSS-WSGKKNKY  503
AN_Q8VY26|_AtCCD8   451 ----------------------------------------GRAMDMCSINPL-YLGQKYRY  470
AN_P74334|_S.ACO    366 ----------------------------------------SRCCEFPVVHPQ-QVGRPYRY  385
AN_Q9NKW9|_DmNinaB  459 RYEEENLVNLVTMEGSQAEAFQGTN-GIQLRPEM------LCDWGCETPRIYYERYMGKNYRY  514
AN_Q6E4P5|_LeCCD1   400 ----------------------------------------ESAVDFPRINEN-YTGRKQRY  419
An_Q8LP16|_PsNCED2  475 ----------------------------------------HMNLEAGMVNKN-KLGRKTQF  494
AN_Q8LP15|_PsNCED3  477 ----------------------------------------QMNLEAGMVNRN-RLGRKTRF  496
AN_Q9AXZ3|_PaNCED1  443 ----------------------------------------HLNLEVGMVNRN-RLGRGPGV  462
AN_Q9AXZ4|_PaNCED3  497 ----------------------------------------PLNLEAGMVNRN-RLGRKTRF  516
AN_C5H805|_LeNCED2  302 ----------------------------------------QVNLEAGMVNRN-KLGRKTQF  321
AN_Q70YP8|_BoLCO    233 ----------------------------------------AENLDFGVIHRG-YSGRKSRY  252
AN_Q84K96|_CsZCD    233 ----------------------------------------AENLDFGVIHRG-YSGRKSRY  252
AN_Q8LP17|_PsCCD1   396 ----------------------------------------ASTVDFPRVNES-YTGRKQRY  415
AN_O06785|_MtCCO    374 ----------------------------------------DRAQEFPRINET-LVGGPHRF  393
AN_Q99NF1|_MmBCO2   379 --------NLSPLSYSSASAVKQGDGEIWCSPENLHHEDLEEEGGIEFPQINYGRFNGKKYSF  433
AN_Q9JJS6|_MmBCO1   374 --------NLVKVSSTTATALKEKDGHVYCQP-------EVLYEGLELPRINYA-YNGKPYRY  420
AN_Q2GA76|_NaCCO    349 ----------------------------------------DTAAEFPRIDDR-FTGQKTRH  368
AN_Q53353|_PpLSD    346 ----------------------------------------DRPDEFPRIDER-YATRAYRH  365
An_O24023|_LeNCED1  478 ----------------------------------------QVNLEAGMVNRN-KLGRKTEY  497
```

```
AN_Q5GN50|_GfCarX   477 ----------------------------------------AIPFEFPSVRPD-REMQEARY  496
cons               1009                                        :   :            1071

AN_Q28175|_BtRPE65  434 AYGLGLNH--F-----------V-PDRLCKLNVKT----------------------------  454
AN_Q16518|_HsRPE65  434 AYGLGLNH--F-----------V-PDRLCKLNVKT----------------------------  454
AN_Q9BYV7|_HsBCO2   481 FYGCGFRH--L-----------V-GDSLIKVDVVN----------------------------  501
AN_Q9HAY6|_HsBCO1   421 VFATGVQW-SP-----------I-PTKIIKYDILT----------------------------  442
AN_O24592|_ZmNCED1  493 AYLAVAEP-WP-----------K-VSGFAKVDLST----------------------------  514
AN_Q45VT7|_ZmCCD1   424 VYCTILDS-IA-----------K-VTGIIKFDLHA----------------------------  445
AN_Q9M9F5|_AtNCED9  550 AYLALTEP-WP-----------K-VSGFAKVDLST----------------------------  571
AN_Q9LRR7|_AtNCED3  493 AYLALAEP-WP-----------K-VSGFAKVDLTT----------------------------  514
AN_Q9C6Z1|_AtNCED5  482 AYLAIAEP-WP-----------K-VSGFAKVDLST----------------------------  503
AN_O49505|_AtNCED2  474 AFLAIAYP-WP-----------K-VSGFAKVDLCT----------------------------  495
AN_Q9LRM7|_AtNCED6  471 VYIAIADP-WP-----------K-CSGIAKVDIQN----------------------------  492
AN_O49675|_AtCCD4   480 VYAAIGDP-MP-----------K-ISGVVKLDVSK----------------------------  501
AN_O65572|_AtCCD1   416 VYGTILDS-IA-----------K-VTGIIKFDLHA----------------------------  437
AN_Q7XJM2|_AtCCD7   504 MYSAASSG-TRS--EL---PHFP-FDMVVKFDLDS----------------------------  531
AN_Q8VY26|_AtCCD8   471 VYACGAQRPCN-----------F-PNALSKVDIVE----------------------------  493
AN_P74334|_S.ACO    386 VYMGAAHH-ST--------GNAP-LQAILKVDLES----------------------------  410
AN_Q9NKW9|_DmNinaB  515 FYAISSDV-DA-----------VNPGTLIKVDVWN----------------------------  537
AN_Q6E4P5|_LeCCD1   420 VYGTTLNS-IA-----------K-VTGVIKFDLHA----------------------------  441
An_Q8LP16|_PsNCED2  495 AYLALAEP-WP-----------K-VSGFAKVDLFS----------------------------  516
AN_Q8LP15|_PsNCED3  497 AYLAVAEP-WP-----------K-VSGFAKVDLES----------------------------  518
AN_Q9AXZ3|_PaNCED1  463 SLSSHCRP-WP-----------K-VSGFAKVDLST----------------------------  484
AN_Q9AXZ4|_PaNCED3  517 AYLAIAEP-WP-----------K-VSGIAKVDLGT----------------------------  538
AN_C5H805|_LeNCED2  322 AYLAIAEP-WP-----------K-VSGFAKVDLST----------------------------  343
AN_Q70YP8|_BoLCO    253 AYLGVGDP-MP-----------K-IRGVVKVDFEL----------------------------  274
AN_Q84K96|_CsZCD    253 AYLGVGDP-MP-----------K-IRGVVKVDFEL----------------------------  274
AN_Q8LP17|_PsCCD1   416 VYGTTLDS-IA-----------K-VTGIIKFDLHA----------------------------  437
AN_O06785|_MtCCO    394 AYTVGIEG-GF--LVGA--GAAL-STPLYKQDCVT----------------------------  422
AN_Q99NF1|_MmBCO2   434 FYGCGFRH--L-----------V-GDSLIKVDVTN----------------------------  454
AN_Q9JJS6|_MmBCO1   421 IFAAEVQW-SP-----------V-PTKILKYDILT----------------------------  442
AN_Q2GA76|_NaCCO    369 GWFLEMDMKRPVELRGGSAGGLL-MNCLFHKDFET----------------------------  402
AN_Q53353|_PpLSD    366 GWMLILDTEKPYEAPGGA-FYAL-TNTLGHIDLAT----------------------------  398
An_O24023|_LeNCED1  498 AYLAIAEP-WP-----------K-VSGFAKVNLFT----------------------------  519
AN_Q5GN50|_GfCarX   497 IYGCSTST-SCFGVALG--RADK-VDLLVKMDAKTLIQRGKKMNATSITGCVDRRSVCEILQE  555
cons               1072                            . : :                         1134

AN_Q28175|_BtRPE65  455 ----KE-------------TWVW-QEPDSYPSEPIFVSHP-D-----ALEEDDGVVLSVVVSP  493
AN_Q16518|_HsRPE65  455 ----KE-------------TWVW-QEPDSYPSEPIFVSHP-D-----ALEEDDGVVLSVVVSP  493
AN_Q9BYV7|_HsBCO2   502 ----KT-------------LKVW-REDGFYPSEPVFVPAP-G-----TNEEDGGVILSVVITP  540
AN_Q9HAY6|_HsBCO1   443 ----KS-------------SLKW-REDDCWPAEPLFVPAP-G-----AKDEDDGVILSAIVST  481
AN_O24592|_ZmNCED1  515 ----GE-------------LTKFEYGEGRFGGEPCFVPMDPAA--AHPRGEDDGYVLTFVHDE  558
AN_Q45VT7|_ZmCCD1   446 ----EPESGVKVLEVGGNVQGIYDLGPGRFGSEAIFVPKHPG----VSGEEDDGYLIFFVHDE  500
AN_Q9M9F5|_AtNCED9  572 ----GE-------------IRKYIYGEGKYGGEPLFLPSG-D------GEEDGGYIMVFVHDE  610
AN_Q9LRR7|_AtNCED3  515 ----GE-------------VKKHLYGDNRYGGEPLFLPGE-G------GEEDEGYILCFVHDE  553
AN_Q9C6Z1|_AtNCED5  504 ----GE-------------VKNHFYGGKKYGGEPFFLPRGLE-----SDGEDDGYIMSFVHDE  544
AN_O49505|_AtNCED2  496 ----GE-------------MKKYIYGGEKYGGEPFFLPGNSGN---GEENEDDGYIFCHVHDE  538
AN_Q9LRM7|_AtNCED6  493 ----GT-------------VSEFNYGPSRFGGEPCFVPEG-E------GEEDKGYVMGFVRDE  531
AN_O49675|_AtCCD4   502 ----GDRD--------DCTVARRMYGSGCYGGEPFFVARDPGN---PEAEEDDGYVVTYVHDE  549
AN_O65572|_AtCCD1   438 ----EAETGKRMLEVGGNIKGIYDLGEGRYGSEAIYVPRE-T------AEEDDGYLIFFVHDE  489
AN_Q7XJM2|_AtCCD7   532 ----NL-------------VRTWSTGARRFVGEPMFVPKNSVE---EGEEEDDGYIVVVEYAV  574
AN_Q8VY26|_AtCCD8   494 ----KK-------------VKNW-HEHGMIPSEPFFVPRP-G-----ATHEDDGVVISIVSEE  532
AN_P74334|_S.ACO    411 ----GT-------------ETLRSFAPHGFAGEPIFVPRPG------GVAEDDGWLLCLIYKA  450
AN_Q9NKW9|_DmNinaB  538 ----KS-------------CLTW-CEENVYPSEPIFVPSP-D-----PKSEDDGVILASMVLG  576
AN_Q6E4P5|_LeCCD1a  442 ----EPETGKSQLEVGGNVQGIFDLGPGRFGSEAVFVPSRPG----TEREEDDGYLIFFVHDE  496
An_Q8LP16|_PsNCED2  517 ----GE-------------VKKYLYGENRFGGEPLFLPNE-D-----SENEDDGYILTFVHDE  556
AN_Q8LP15|_PsNCED3  519 ----GE-------------VKKHSYGDGRFGGEPFFLPTRRSGRCEYENEEDGGYIMALVHDE  564
AN_Q9AXZ3|_PaNCED1  485 ----GE-------------VTKFIYGEQCYGGEPYFVSRDPV------APEDDGYVLSFMHDE  524
AN_Q9AXZ4|_PaNCED3  539 ----GE-------------VNRFVYGERQFGGEPYFIPREPS----TSGREDDGYVVSFMHDE  580
AN_C5H805|_LeNCED2  344 ----GE-------------IKKHIYGDKRYGGEPLFLPRNVN-----SEKEDDGYILAFCHDE  384
AN_Q70YP8|_BoLCO    275 ----AGRG--------ECVVARREFGVGCFGGEPFFVPASSKK---SGGEEDDGYVVSYLHDE  322
AN_Q84K96|_CsZCD    275 ----AGRG--------ECVVARREFGVGCFGGEPFFVPASSKK---SGGEEDDGYVVSYLHDE  322
AN_Q8LP17|_PsCCD1   438 ----EPDSGKTKLEVGGNVQGLYDLGPGRFGSEAVYVPRVPG----TDSEEDDGYLIFFVHDE  492
AN_O06785|_MtCCO    423 ----GS-------------STVASLDPDLLIGEMVFVPNP-S-----ARAEDDGILMGYGWHR  462
AN_Q99NF1|_MmBCO2   455 ----KT-------------LRVW-REEGFYPSEPVFVPVP-G-----ADEEDSGVILSVVITP  493
AN_Q9JJS6|_MmBCO1   443 ----KS-------------SLKW-SEESCWPAEPLFVPTP-G-----AKDEDDGVILSAIVST  481
AN_Q2GA76|_NaCCO    403 ----GR-------------EQHWWCGPVSSLQEPCFVPRA-K-----DAPEGDGWIVQVCNRL  442
AN_Q53353|_PpLSD    399 ----GK-------------SSSWWAGPRCAIQEPCFIPRS-P-----DAPEGDGYVIALVDDH  438
```

```
An_Q24023|_LeNCED1  520 ----GE------------VEKFIYGDNKYGGEPLFLPRDPN-----SKEEDDGYILAFVHDE  560
AN_Q5GN50|_GfCarX   556 QRKDDP------------IYIFRLPPNHYAQEPRFVPRAC------STEEDDGYLLFYVFDE  599
cons               1135                                *  ::.           *. * :.       1197

AN_Q28175|_BtRPE65  494 GA----------GQKPAYLLILNAKDLSE--VARAE--VEINI-------PVTFHGLFKK---  532
AN_Q16518|_HsRPE65  494 GA----------GQKPAYLLILNAKDLSE--VARAE--VEINI-------PVTFHGLFKK---  532
AN_Q9BYV7|_HsBCO2   541 NQ----------NE-SNFILVLDAKNFEE--LGRAE--VPVQM-------PYGFHGTFIP---  578
AN_Q9HAY6|_HsBCO1   482 DP----------QK-LPFLLILDAKSFTE--LARAS--VDVDM-------HMDLHGLFITDMD  522
AN_O24592|_ZmNCED1  559 R----------AGT--SELLVVNAADMRL--EATVQ--LPSRV-------PFGFHGTFITGQE  598
AN_Q45VT7|_ZmCCD1   501 N----------TGK--SEVNVIDAKTMSADPVAVVE--LPNRV-------PYGFHAFFVTEDQ  542
AN_Q9M9F5|_AtNCED9  611 E----------KVK--SELQLINAVNMKL--EATVT--LPSRV-------PYGFHGTFISKED  650
AN_Q9LRR7|_AtNCED3  554 K----------TWK--SELQIVNAVSLEV--EATVK--LPSRV-------PYGFHGTFIGADD  593
AN_Q9C6Z1|_AtNCED5  545 E----------SWE--SELHIVNAVTLEL--EATVK--LPSRV-------PYGFHGTFVNSAD  584
AN_O49505|_AtNCED2  539 E----------TKT--SELQIINAVNLKL--EATIK--LPSRV-------PYGFHGTFVDSNE  578
AN_Q9LRM7|_AtNCED6  532 E----------KDE--SEFVVVDATDMKQ--VAAVR--LPERV-------PYGFHGTFVSENQ  571
AN_O49675|_AtCCD4   550 V----------TGE--SKFLVMDAKSPELEIVAAVR--LPRRV-------PYGFHGLFVKESD  591
AN_O65572|_AtCCD1   490 N----------TGK--SCVTVIDAKTMSAEPVAVVE--LPHRV-------PYGFHALFVTEEQ  531
AN_Q7XJM2|_AtCCD7   575 S----------VER--CYLVILDAKKIGE-SDAVVS--SVNKVYIAKINYIICVHSFYFDRNI  622
AN_Q8VY26|_AtCCD8   533 N----------G-G--SFAILLDGSSFEE--IARAK--FPYGL-------PYGLHGCWIPK--  569
AN_P74334|_S.ACO    451 D----------LHR--SELVILDAQDITAPAIATLK--LKHHI-------PYPLHGSWAQ---  489
AN_Q9NKW9|_DmNinaB  577 GL----------NDRYVGLIVLCAKTMTE--LGRCDFHTNGPV-------PKCLHGWFAPN--  618
AN_Q6E4P5|_LeCCD1   497 N----------TGK--SAVNVIDAKTMSAEPVAVVE--LPKRV-------PYGFHAFFVTEEQ  538
An_Q8LP16|_PsNCED2  557 K----------EWK--SEIQIVNAVNLKL--EACIP--LPSRV-------PYGFHGTFIHSKE  596
AN_Q8LP15|_PsNCED3  565 R----------RCK--SELQIVNAVNLEV--EATVK--LPSRV-------PYGFHGTFVEAKD  604
AN_Q9AXZ3|_PaNCED1  525 K----------TAR--SELLIVNAITMQL--EASVK--LPSRV-------PYGFHGTFISSKD  564
AN_Q9AXZ4|_PaNCED3  581 K----------TSR--SELLILNAMNMRL--EASVM--LPSRV-------PYGFHGTFISSRD  620
AN_C5H805|_LeNCED2  385 K----------TWK--SELQIVNAMTLEL--EATVK--LPSRV-------PYGFHGTFISSKD  424
AN_Q70YP8|_BoLCO    323 G----------KGE--SSFVVMDARSAQLEILAEVV--LPRRV-------PYGFHGLFVTDKD  364
AN_Q84K96|_CsZCD    323 G----------KGE--SSFVVMDARSPELEILAEVV--LPRRV-------PYGFHGLFVTEAE  364
AN_Q8LP17|_PsCCD1   493 N----------TGK--SFVHVIDAKRMSAEPVAVVE--LPQRV-------PYGFHAFFVTEDQ  534
AN_O06785|_MtCCO    463 G----------RDE--GQLLLLDAQTLES--IATVH--LPQRV-------PMGFHGNWAP---  499
AN_Q99NF1|_MmBCO2   494 NQ----------SE-SNFLLVLDAKSFTE--LGRAE--VPVQM-------PYGFHGTFVP---  531
AN_Q9JJS6|_MmBCO1   482 DP----------QK-LPFLLILDAKSFTE--LARAS--VDADM-------HLDLHGLFIPDAD  522
AN_Q2GA76|_NaCCO    443 E----------EQR--SDLLIFDALDIEKGPVATVN--IPIRL-------RFGLHGNWANADE  484
AN_Q53353|_PpLSD    439 V----------ANY--SDLAIFDAQHVDQGPIARAK--LPVRI-------RQGLHGNWADASR  480
An_Q24023|_LeNCED1  561 K----------EWK--SELQIVNAMSLKL--EATVK--LPSRV-------PYGFHGTFINAND  600
AN_Q5GN50|_GfCarX   600 SQLLPSGDCPPSAT--SELWILDAKNMRD-VVAKVR--LPQRV-------PYGLHGTWFSSQD  650
cons               1198                    :. .       .        :          .*. :      1260

AN_Q28175|_BtRPE65  533 ------------------------------------------------------------  532
AN_Q16518|_HsRPE65  533 ------------------------------------------------------------  532
AN_Q9BYV7|_HsBCO2   579 ------------------------------------------------------------  578
AN_Q9HAY6|_HsBCO1   523 WDTKKQAASEE--------------QRDR----ASDCHGAP-------------------  545
AN_O24592|_ZmNCED1  599 LE---------------------------------------AQ-----------------  602
AN_Q45VT7|_ZmCCD1   543 LA---------------------------------------RQA----------------  547
AN_Q9M9F5|_AtNCED9  651 LS---------------------------------------KQ-----------------  654
AN_Q9LRR7|_AtNCED3  594 LA---------------------------------------KQ-----------------  597
AN_Q9C6Z1|_AtNCED5  585 ML---------------------------------------N------------------  587
AN_O49505|_AtNCED2  579 LV---------------------------------------D------------------  581
AN_Q9LRM7|_AtNCED6  572 LK---------------------------------------EQ-----------------  575
AN_O49675|_AtCCD4   592 L----------------------------------------N------------------  593
AN_O65572|_AtCCD1   532 LQ---------------------------------------EQ-----------------  535
AN_Q7XJM2|_AtCCD7   623 AF-----------------------------------------------------------  624
AN_Q8VY26|_AtCCD8   570 ------------------------------------------------------------  569
AN_P74334|_S.ACO    490 ------------------------------------------------------------  489
AN_Q9NKW9|_DmNinaB  619 ------------------------------------------------------------  618
AN_Q6E4P5|_LeCCD1a  539 IQ---------------------------------------EQ-----------------  542
An_Q8LP16|_PsNCED2  597 LE---------------------------------------K------------------  599
AN_Q8LP15|_PsNCED3  605 LT---------------------------------------------------------  606
AN_Q9AXZ3|_PaNCED1  565 LA---------------------------------------N------------------  567
AN_Q9AXZ4|_PaNCED3  621 LA---------------------------------------K------------------  623
AN_C5H805|_LeNCED2  425 LQ---------------------------------------N------------------  427
AN_Q70YP8|_BoLCO    365 LL---------------------------------------N------------------  367
AN_Q84K96|_CsZCD    365 LL---------------------------------------S------------------  367
AN_Q8LP17|_PsCCD1   535 LQ---------------------------------------EQ-----------------  538
AN_O06785|_MtCCO    500 ------------------------------------------------------------  499
AN_Q99NF1|_MmBCO2   532 ------------------------------------------------------------  531
AN_Q9JJS6|_MmBCO1   523 WNAVKQTPAETQEVENSDHPTDPTAPELSHSENDFTAGHGGS-------------------  564
AN_Q2GA76|_NaCCO    485 IGLAE----------------------------------KVL-------------------  492
```

```
AN_Q53353|_PpLSD    481 LA-------------------------------------V----------------------  483
An_O24023|_LeNCED1  601 LA-------------------------------------N----------------------  603
AN_Q5GN50|_GfCarX   651 IE-------------------------------------SQRSVESLRSLEVVQRKKEEWV  674
cons               1261                                                              1323

AN_Q28175|_BtRPE65  533 --------------------S  533
AN_Q16518|_HsRPE65  533 --------------------S  533
AN_Q9BYV7|_HsBCO2   579 --------------------I  579
AN_Q9HAY6|_HsBCO1   546 --------------------LT  547
AN_O24592|_ZmNCED1  603 --------------------AA  604
AN_Q45VT7|_ZmCCD1   548 -------------------EGQ  550
AN_Q9M9F5|_AtNCED9  655 -------------------ALC  657
AN_Q9LRR7|_AtNCED3  598 --------------------VV  599
AN_Q9C6Z1|_AtNCED5  588 --------------------QA  589
AN_O49505|_AtNCED2  582 --------------------QL  583
AN_Q9LRM7|_AtNCED6  576 --------------------VF  577
AN_O49675|_AtCCD4   594 --------------------KL  595
AN_O65572|_AtCCD1   536 -------------------TLI  538
AN_Q7XJM2|_AtCCD7   625 -----------------HLSHK  629
AN_Q8VY26|_AtCCD8   570 --------------------D  570
AN_P74334|_S.ACO    490 --------------------T  490
AN_Q9NKW9|_DmNinaB  619 --------------------AI  620
AN_Q6E4P5|_LeCCD1a  543 -------------------AKL  545
An_Q8LP16|_PsNCED2  600 --------------------QE  601
AN_Q8LP15|_PsNCED3  607 --------------------FQ  608
AN_Q9AXZ3|_PaNCED1  568 --------------------QA  569
AN_Q9AXZ4|_PaNCED3  624 --------------------QA  625
AN_C5H805|_LeNCED2  428 --------------------QV  429
AN_Q70YP8|_BoLCO    368 --------------------QA  369
AN_Q84K96|_CsZCD    368 --------------------QQ  369
AN_Q8LP17|_PsCCD1   539 -------------------AKF  541
AN_O06785|_MtCCO    500 --------------------TT  501
AN_Q99NF1|_MmBCO2   532 --------------------I  532
AN_Q9JJS6|_MmBCO1   565 --------------------SL  566
AN_Q2GA76|_NaCCO    493 --------------------AA  494
AN_Q53353|_PpLSD    484 --------------------AA  485
An_O24023|_LeNCED1  604 --------------------QA  605
AN_Q5GN50|_GfCarX   675 NSGGQIRKSWMVLREKLEKAVG  696
cons               1324                       1345
```

10.2 *Zea mays* NCED Sequences

Sequences of the GST-tagged *Z. mays* NCEDs 1, 2, 3A, 3B and 9. Main sequence is shown in red. GST tag is shown in black.

10.2.1 ZmNCED1

Accession number for ZmNCED1: ZMU95953.

```
258 ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTT
  1  M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L

318 TTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAA
 21  L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K

378 TGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT
 41  W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D

438 GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAAC
 61  G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N

498 ATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTG
 81  M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L

558 GATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTT
```

```
 101  D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V

 618 GATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA
 121  D  F  L  S  K  L  P  E  M  L  K  M  F  E  D  R  L  C  H  K

 678 ACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGAT
 141  T  Y  L  N  G  D  H  V  T  H  P  D  F  M  L  Y  D  A  L  D

 738 GTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAA
 161  V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K

 798 AAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCA
 181  K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A

 858 TGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
 201  W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D

 918 CTGGTTCCGCGTGGATCCGCCTCCAATTCCGTCAGGTTCTCGCCGCGCGCCGTCAGCTCC
 221  L  V  P  R  G  S  A  S  N  S  V  R  F  S  P  R  A  V  S  S

 978 GTGCCGCCCGCCGAGTGCCTCCAAGCGCCGTTCCACAAGCCCGTCGCCGACCTGCCGGCG
 241  V  P  P  A  E  C  L  Q  A  P  F  H  K  P  V  A  D  L  P  A

1038 CCGTCCAGGAAGCCCGCCGCCATTGCCGTCCCAGGGCACTCCGCGGCGCCGAGGAAAGCG
 261  P  S  R  K  P  A  A  I  A  V  P  G  H  S  A  A  P  R  K  A

1098 GACGGCGGCAAGAAGCAGCTCAACTTGTTCCAGCGCGCCGCGGCGGCCGCGCTCGACGCG
 281  D  G  G  K  K  Q  L  N  L  F  Q  R  A  A  A  A  A  L  D  A

1158 TTCGAGGAAGGGTTCGTGGCCAACGTCCTGGAGCGGCCCCACGGGCTGCCCAGCACGGCC
 301  F  E  E  G  F  V  A  N  V  L  E  R  P  H  G  L  P  S  T  A

1218 GACCCGGCCGTGCAGATCGCCGGCAACTTCGCGCCCGTCGGGGAGAGCCCGCCCGTGCAC
 321  D  P  A  V  Q  I  A  G  N  F  A  P  V  G  E  S  P  P  V  H

1278 GAGCTCCCCGTCTCCGGCCGCATCCCGCCCTTCATCGACGGGGTCTACGCGCGCAACGGC
 341  E  L  P  V  S  G  R  I  P  P  F  I  D  G  V  Y  A  R  N  G

1338 GCCAACCCCTGCTTCGACCCCGTCGCCGGGCACCACCTCTTCGACGGCGACGGCATGGTG
 361  A  N  P  C  F  D  P  V  A  G  H  H  L  F  D  G  D  G  M  V

1398 CACGCGCTGCGGATACGCAACGGCGCCGCCGAGTCCTACGCCTGCCGCTTCACGGAGACC
 381  H  A  L  R  I  R  N  G  A  A  E  S  Y  A  C  R  F  T  E  T

1458 GCGCGCCTGCGCCAGGAGCGCGCGATCGGCCGCCCCGTCTTCCCCAAGGCCATTGGCGAG
 401  A  R  L  R  Q  E  R  A  I  G  R  P  V  F  P  K  A  I  G  E

1518 CTGCACGGGCACTCCGGGATCGCGCGCCTCGCCCTGTTCTACGCGCGCGCCGCGTGCGGC
 421  L  H  G  H  S  G  I  A  R  L  A  L  F  Y  A  R  A  A  C  G

1578 CTCGTGGACCCCTCGGCCGGCACCGGCGTGGCCAACGCCGGCCTCGTCTACTTCAACGGC
 441  L  V  D  P  S  A  G  T  G  V  A  N  A  G  L  V  Y  F  N  G

1638 CGCCTGCTCGCCATGTCCGAGGACGACCTCCCCTACCACGTCCGCGTGGCGGACGACGGC
 461  R  L  L  A  M  S  E  D  D  L  P  Y  H  V  R  V  A  D  D  G

1698 GACCTCGAGACCGTCGGCCGCTACGACTTCGACGGGCAGCTCGGCTGCGCCATGATCGCG
 481  D  L  E  T  V  G  R  Y  D  F  D  G  Q  L  G  C  A  M  I  A

1758 CACCCCAAGCTGGACCCGGCCACCGGGGAGCTCCACGCGCTCAGCTACGACGTCATCAAG
 501  H  P  K  L  D  P  A  T  G  E  L  H  A  L  S  Y  D  V  I  K

1818 AGGCCGTACCTCAAGTACTTCTACTTCAGGCCCGACGGCACCAAGTCCGACGACGTGGAG
 521  R  P  Y  L  K  Y  F  Y  F  R  P  D  G  T  K  S  D  D  V  E

1878 ATCCCGCTGGAGCAGCCTACGATGATCCACGACTTCGCCATCACCGAGAACTTCGTGGTT
 541  I  P  L  E  Q  P  T  M  I  H  D  F  A  I  T  E  N  F  V  V

1938 GTGCCCGACCACCAGGTGGTGTTCAAGCTCCAGGAGATGCTGCGCGGCGGGTCGCCCGTG
```

```
 561  V  P  D  H  Q  V  V  F  K  L  Q  E  M  L  R  G  G  S  P  V

1998 GTGCTGGACAAGGAGAAGACGTCGCGGTTCGGCGTGCTCCCCAAGCACGCCGCGGACGCG
 581  V  L  D  K  E  K  T  S  R  F  G  V  L  P  K  H  A  A  D  A

2058 TCGGAGATGGCGTGGGTGGACGTGCCGGACTGCTTCTGCTTCCACCTGTGGAACGCGTGG
 601  S  E  M  A  W  V  D  V  P  D  C  F  C  F  H  L  W  N  A  W

2118 GAGGACGAGGCGACGGGCGAGGTGGTGGTGATCGGCTCCTGCATGACCCCCGCCGACTCC
 621  E  D  E  A  T  G  E  V  V  V  I  G  S  C  M  T  P  A  D  S

2178 ATCTTCAACGAGTCCGACGAGCGCCTGGAGAGCGTGCTGACGGAGATCCGCCTGGACGCG
 641  I  F  N  E  S  D  E  R  L  E  S  V  L  T  E  I  R  L  D  A

2238 CGCACGGGCCGGTCCACGCGCCGCGCCGTCCTGCCGCCGTCGCAGCAGGTGAACCTGGAG
 661  R  T  G  R  S  T  R  R  A  V  L  P  P  S  Q  Q  V  N  L  E

2298 GTGGGCATGGTGAACCGCAACCTGCTGGGCCGCGAGACCCGGTACGCGTACCTCGCGGTG
 681  V  G  M  V  N  R  N  L  L  G  R  E  T  R  Y  A  Y  L  A  V

2358 GCGGAGCCGTGGCCCAAGGTGTCGGGCTTCGCCAAGGTGGACCTGTCCACGGGCGAGCTC
 701  A  E  P  W  P  K  V  S  G  F  A  K  V  D  L  S  T  G  E  L

2418 ACCAAGTTCGAGTACGGCGAGGGCCGGTTCGGCGGCGAGCCATGCTTCGTGCCCATGGAC
 721  T  K  F  E  Y  G  E  G  R  F  G  G  E  P  C  F  V  P  M  D

2478 CCGGCCGCGGCCCACCCGCGCGGCGAGGACGACGGGTACGTGCTCACCTTCGTCCACGAC
 741  P  A  A  A  H  P  R  G  E  D  D  G  Y  V  L  T  F  V  H  D

2538 GAGCGCGCCGGCACGTCGGAGCTACTTGTGGTCAATGCCGCCGACATGCGGCTGGAGGCC
 761  E  R  A  G  T  S  E  L  L  V  V  N  A  A  D  M  R  L  E  A

2598 ACGGTCCAGCTGCCGTCCCGCGTGCCCTTCGGCTTCCACGGCACCTTCATCACGGGCCAG
 781  T  V  Q  L  P  S  R  V  P  F  G  F  H  G  T  F  I  T  G  Q

2658 GAGCTCGAGGCCCAGGCGGCCTGA
 801  E  L  E  A  Q  A  A  *
```

10.2.2 ZmNCED2

```
258 ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTT
  1  M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L

318 TTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAA
 21  L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K

378 TGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT
 41  W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D

438 GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAAC
 61  G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N

498 ATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTG
 81  M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L

558 GATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTT
101  D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V

618 GATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA
121  D  F  L  S  K  L  P  E  M  L  K  M  F  E  D  R  L  C  H  K

678 ACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGAT
141  T  Y  L  N  G  D  H  V  T  H  P  D  F  M  L  Y  D  A  L  D

738 GTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAA
161  V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K
```

```
 798 AAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCA
 181  K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A

 918 TGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
 201  W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D

 978 CTGGTTCCGCGTGGATCCGCAGCGGCAGCGGCCAACTCGGTCTTCAGCGCGCCGCCCGCA
 221  L  V  P  R  G  S  A  A  A  A  A  N  S  V  F  S  A  P  P  A

1038 GCCGTGCGCTACGCCGCGCCGCCCAAGCCGGCGGCGCCCATTGCGCCAGTGCCCGCTCCG
 241  A  V  R  Y  A  A  P  P  K  P  A  A  P  I  A  P  V  P  A  P

1098 GCGCCGGCGGCGGGCGGCGAGAAAGGCCTGAGCTTCCTGCAGCGCGCGGCGTCGGCGGCG
 261  A  P  A  A  G  G  E  K  G  L  S  F  L  Q  R  A  A  S  A  A

1158 CTGGACGCGTTCGAATCCGGCGTGATCGCGGGCCTCCTGGAGCGTCCCCGCGCGCTGCCC
 281  L  D  A  F  E  S  G  V  I  A  G  L  L  E  R  P  R  A  L  P

1218 CGGTCGGCGGACCCGGCCGTGCAGATCGCCGGCAACTTCGCCCCCGTCGGCGAGCAGCCG
 301  R  S  A  D  P  A  V  Q  I  A  G  N  F  A  P  V  G  E  Q  P

1278 CCCGTGCGGTCGCTGCCCGTGTCGGGCCGCATCCCGCCCTTCATCTCCGGCGTGTACGCG
 321  P  V  R  S  L  P  V  S  G  R  I  P  P  F  I  S  G  V  Y  A

1338 CGCAACGGCGCCAACCCCTGCTTCGACCCCGCCGCGGGGCACCACCTGTTCGACGGCGAC
 341  R  N  G  A  N  P  C  F  D  P  A  A  G  H  H  L  F  D  G  D

1398 GGCATGGTGCACGCGGTCCGCATCCGCAACGGCGCCGCGGAGTCGTACGCGTGCCGCTTC
 361  G  M  V  H  A  V  R  I  R  N  G  A  A  E  S  Y  A  C  R  F

1458 ACCGAGACGGCGCGGCTCCGGCAGGAGCGCGCCCTGGGGCGGGCCGTGTTCCCCAAGGCC
 381  T  E  T  A  R  L  R  Q  E  R  A  L  G  R  A  V  F  P  K  A

1518 ATCGGCGAGCTGCACGGCCACTCCGGGATCGCGCGCCTGGCGCTCTTCTACGCGCGCGGC
 401  I  G  E  L  H  G  H  S  G  I  A  R  L  A  L  F  Y  A  R  G

1578 CTCTGCGGCCTCGTCGACGCGTCCCGCGGCACGGGCGTCGCCAACGCCGGCCTCGTCTAC
 421  L  C  G  L  V  D  A  S  R  G  T  G  V  A  N  A  G  L  V  Y

1638 TTCAACGGCCGCCTGCTGGCCATGTCCGAGGACGACCTACCGTACCAGGTGCGCGTGACC
 441  F  N  G  R  L  L  A  M  S  E  D  D  L  P  Y  Q  V  R  V  T

1698 GCCGACGGGGACCTCCGCACCGTCGGGCGCTACGACTTCGAAGGGCAGCTCGGGTGCGCC
 461  A  D  G  D  L  R  T  V  G  R  Y  D  F  E  G  Q  L  G  C  A

1758 AGCATGATCGCGCACCCCAAGCTGGACCCGGCCTCCGGCGACCTCTTCGCGCTCAGCTAC
 481  S  M  I  A  H  P  K  L  D  P  A  S  G  D  L  F  A  L  S  Y

1818 GACGTCATCCGGCGGCCCTACCTGAGGTACTTCTACTTCCGGCCCGACGGCACCAAGTCC
 501  D  V  I  R  R  P  Y  L  R  Y  F  Y  F  R  P  D  G  T  K  S

1878 GACGACGTGGAGATCCCGCTGGAGCAGCCGACCATGATCCACGACTTCGCCATCACGGAG
 521  D  D  V  E  I  P  L  E  Q  P  T  M  I  H  D  F  A  I  T  E

1938 CGGTTCGTGGTGGTGCCCGACCACCAGGTGGTGTTCAAGCTCGGCGAGATGCTCCGCGGC
 541  R  F  V  V  V  P  D  H  Q  V  V  F  K  L  G  E  M  L  R  G

1998 GGGTCCCCCGTGGTGCTGGACCAGGGCAAGACGTCCCGCTTCGGCGTGCTGCCCAAGTAC
 561  G  S  P  V  V  L  D  Q  G  K  T  S  R  F  G  V  L  P  K  Y

2058 GCGCGCGACGCGTCGGAGATGGTGTGGGTGGACGTGCCGGACTGCTTCTGCTTCCACCTG
 581  A  R  D  A  S  E  M  V  W  V  D  V  P  D  C  F  C  F  H  L

2118 TGGAACGCGTGGGAGGAGCCCGGCACGGGGGAGGTGGTGGTGGTCGGGTCCTGCATGACC
 601  W  N  A  W  E  E  P  G  T  G  E  V  V  V  V  G  S  C  M  T

2178 CCCGCCGACTCCATCTTCAACGACGACGGCGGCGAGGACGGGCGGCTCCAGAGCGTGCTG
```

```
 621  P  A  D  S  I  F  N  D  D  G  G  E  D  G  R  L  Q  S  V  L

2238 ACCGAGATCCGGCTGGACACGCGCACGGGCGCGTCCACGCGCCGCGCCGTGCTGCCTCCG
 641  T  E  I  R  L  D  T  R  T  G  A  S  T  R  R  A  V  L  P  P

2298 TCGGCGCAGGTGAACCTGGAGGTCGGGATGGTGAACCGCAACATGGTGGGGCGGAGGACG
 661  S  A  Q  V  N  L  E  V  G  M  V  N  R  N  M  V  G  R  R  T

2358 CGGTACGCGTACCTGGCCGTGGCGGAGCCGTGGCCCAAGGTGTCCGGGTTCGCCAAGGTG
 681  R  Y  A  Y  L  A  V  A  E  P  W  P  K  V  S  G  F  A  K  V

2418 GACCTGGACACGGGCGACCTCGTCCGGTTCGACTACGGCGACGGCCGCTTCGGCGGGGAG
 701  D  L  D  T  G  D  L  V  R  F  D  Y  G  D  G  R  F  G  G  E

2478 CCCTGCTTCGTGCCCACCGAGGGCGCGCCGGCGCGCGGCGAGGACGACGGCTACATCCTG
 721  P  C  F  V  P  T  E  G  A  P  A  R  G  E  D  D  G  Y  I  L

2538 TCCCTCGTCCGCGACGAGCGGGCGGGCACGTCGGAGCTCCTGGTAGTGAACGCCGCCGAC
 741  S  L  V  R  D  E  R  A  G  T  S  E  L  L  V  V  N  A  A  D

2598 ATGCGGCTGGAGGCCACGGTCCAGCTACCCTCCCGCGTGCCCTATGGCTTCCACGGCACC
 761  M  R  L  E  A  T  V  Q  L  P  S  R  V  P  Y  G  F  H  G  T

2658 TTCATCGGCGAACGGGAGCTAGAGGCCCAGGCCTGA
 781  F  I  G  E  R  E  L  E  A  Q  A  *
```

10.2.3 ZmNCED3A

```
515 ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTT
   1  M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L

 575 TTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAA
  21  L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K

 635 TGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT
  41  W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D

 695 GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAAC
  61  G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N

 755 ATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTG
  81  M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L

 815 GATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTT
 101  D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V

 875 GATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA
 121  D  F  L  S  K  L  P  E  M  L  K  M  F  E  D  R  L  C  H  K

 935 ACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGAT
 141  T  Y  L  N  G  D  H  V  T  H  P  D  F  M  L  Y  D  A  L  D

 995 GTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAA
 161  V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K

1055 AAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCA
 181  K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A

1115 TGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
 201  W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D

1175 CTGGTTCCGCGTGGATCCGCCGCCGCTCCCAAGTGGAACCCGCTCCAGCGGCTCGCCGCG
 221  L  V  P  R  G  S  A  A  A  P  K  W  N  P  L  Q  R  L  A  A

1235 GCGGCGCTCGACGCGCTGGAGGAGGGGCTCGTCGCGGGCGTCCTCGAGCGCGCGCACCCG
 241  A  A  L  D  A  L  E  E  G  L  V  A  G  V  L  E  R  A  H  P
```

```
1295 CTGCCACGGACCGCCGACCCGGCCGTCCAGATCGCCGGCAACTACGCGCCCGTCGGGGAG
 261  L  P  R  T  A  D  P  A  V  Q  I  A  G  N  Y  A  P  V  G  E

1355 CGCCCGCCCACCGGCGAGCTGCCCGTGTCCGGCCGCGTCCCGGCGTGCCTCGACGGCGTG
 281  R  P  P  T  G  E  L  P  V  S  G  R  V  P  A  C  L  D  G  V

1415 TACGTCCGCAACGGCGCCAACCCGCTCCACGCGCCGCGCGCGGGGCACCACCTCTTCGAC
 301  Y  V  R  N  G  A  N  P  L  H  A  P  R  A  G  H  H  L  F  D

1475 GGCGACGGCATGCTGCACGCCGTGCGGCTCCGCGGGGGCCGCGCCGAGTCGTACGCGTGC
 321  G  D  G  M  L  H  A  V  R  L  R  G  G  R  A  E  S  Y  A  C

1535 CGGTTCACGGAGACGGCGCGGCTCAGGCAGGAGCGCGCCATCGGCAGGGCCGTGTTCCCC
 341  R  F  T  E  T  A  R  L  R  Q  E  R  A  I  G  R  A  V  F  P

1595 AAGGCCATCGGCGAGCTGCACGGCCACTCGGGCGTGGCGCGCCTGCTCCTCTTCGGCGCC
 361  K  A  I  G  E  L  H  G  H  S  G  V  A  R  L  L  L  F  G  A

1655 CGCTCCCTCTGCGGCGTGCTCGACGCGTCCCAGGGGATCGGGGTCGCCAACGCCGGCCTG
 381  R  S  L  C  G  V  L  D  A  S  Q  G  I  G  V  A  N  A  G  L

1715 GTGTACCACAACAACCGCCTCCTCGCCATGTCGGAGGACGACCTCCCCTACCACGTGCGC
 401  V  Y  H  N  N  R  L  L  A  M  S  E  D  D  L  P  Y  H  V  R

1775 GTCACCGCCGACGGCGACCTCGAGACCGCCGGGCGCTACGACTTCGGCGGCCAGCTCGAC
 421  V  T  A  D  G  D  L  E  T  A  G  R  Y  D  F  G  G  Q  L  D

1835 ACCGCCATGATCGCGCACCCCAAGCTGGACCCGGCCACCGGCGAGCTCTTCGCGCTCAGC
 441  T  A  M  I  A  H  P  K  L  D  P  A  T  G  E  L  F  A  L  S

1895 TACAATGTCGTGTCCAAGCCGTTCCTCAAGTACTTCTACTTCACCGCCGACGGCCGCAAG
 461  Y  N  V  V  S  K  P  F  L  K  Y  F  Y  F  T  A  D  G  R  K

1955 TCCCCCGACGTCGAGATCCCCGTGGACGCGCCCACCATGATGCACGACTTCGCCGTCACC
 481  S  P  D  V  E  I  P  V  D  A  P  T  M  M  H  D  F  A  V  T

2015 GAGAACCACGCCATCATCCCGGACCAGCAGATCGTGTTCAAGCTCCAGGAGATGGTGCTG
 501  E  N  H  A  I  I  P  D  Q  Q  I  V  F  K  L  Q  E  M  V  L

2075 GGCGGCTCCCCCGTGGTGTACGACAGGAACAAGACCGCGCGGTTCGGGGTGCTGCCGAAG
 521  G  G  S  P  V  V  Y  D  R  N  K  T  A  R  F  G  V  L  P  K

2135 CGCGCCACCGACGCGTCGCGGCTGCGGTGGGTGGACGTCCCCGACTGCTTCTGCTTCCAC
 541  R  A  T  D  A  S  R  L  R  W  V  D  V  P  D  C  F  C  F  H

2195 CTCTGGAACGCGTGGGAGGACGAGGCCACGGGCGAGGTCGTGGTGATCGGGTCCTGCATG
 561  L  W  N  A  W  E  D  E  A  T  G  E  V  V  V  I  G  S  C  M

2255 ACCCCCGCGGACGCCGTGTTCAACGAGTCCGCCGCCGGCGAGGAGAGCTTCCGCAGCGTG
 581  T  P  A  D  A  V  F  N  E  S  A  A  G  E  E  S  F  R  S  V

2315 CTGTCGGAGATCCGGCTGGACCCGCGCACCGGCACGTCCTCGCGGCGCGCGGTCCTGAGC
 601  L  S  E  I  R  L  D  P  R  T  G  T  S  S  R  R  A  V  L  S

2375 GACGCCGACCAGGTGAACCTGGAGGCCGGCATGGTGAACCGGCAGCTGCTGGGCAGGAGG
 621  D  A  D  Q  V  N  L  E  A  G  M  V  N  R  Q  L  L  G  R  R

2435 ACGCGCTACGCCTACCTCGCCATCGCCGAGCCGTGGCCCAAGGTGTCGGGCTTCGCCAAG
 641  T  R  Y  A  Y  L  A  I  A  E  P  W  P  K  V  S  G  F  A  K

2495 GTGGACCTCGAGGCCGGCACCGTCGACAAGTTCATCTACGGCGACGGCCGGTACGGCGGC
 661  V  D  L  E  A  G  T  V  D  K  F  I  Y  G  D  G  R  Y  G  G

2555 GAGCCCTGCTTCGTGCCGCGCCCCGACGCCCCCGCGGGCGCCGCGGAGGACGACGGCTAC
 681  E  P  C  F  V  P  R  P  D  A  P  A  G  A  A  E  D  D  G  Y

2615 GTGCTGTGCTACGTGCACGACGAGGCCCGCGGTGCGTCGGAAATGCTCGTCGTCAACGCC
 701  V  L  C  Y  V  H  D  E  A  R  G  A  S  E  M  L  V  V  N  A
```

```
2675 CGCGACATGCGGGCCGAGGCCGCCGTCAAGCTGCCGGGCCGCGTCCCGTACGGGTTGCAC
 721  R  D  M  R  A  E  A  A  V  K  L  P  G  R  V  P  Y  G  L  H

2735 GGCACGTTCATCGCCGGCAAGGAGCTGCAGCGGCAGGCCTAG
 741  G  T  F  I  A  G  K  E  L  Q  R  Q  A  *
```

10.2.4 ZmNCED3B

```
 258 ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTT
   1  M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L

 318 TTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAA
  21  L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K

 378 TGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT
  41  W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D

 438 GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAAC
  61  G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N

 498 ATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTG
  81  M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L

 558 GATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTT
 101  D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V

 618 GATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA
 121  D  F  L  S  K  L  P  E  M  L  K  M  F  E  D  R  L  C  H  K

 678 ACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGAT
 141  T  Y  L  N  G  D  H  V  T  H  P  D  F  M  L  Y  D  A  L  D

 738 GTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAA
 161  V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K

 798 AAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCA
 181  K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A

 858 TGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
 201  W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D

 918 CTGGTTCCGCGTGGATCCGCGGCCGCCAAGTGGAACCCGCTCCAGCGGCTCGCCGCGGCG
 221  L  V  P  R  G  S  A  A  A  K  W  N  P  L  Q  R  L  A  A  A

 978 GCGCTCGACGCGCTGGAGGAGGGGCTCGTGGCGGGCGTCCTGGAGCGCGCGCACCCGCTG
 241  A  L  D  A  L  E  E  G  L  V  A  G  V  L  E  R  A  H  P  L

1038 CCACGGACCGCCGACCCGGCCGTCCAGATCGCCGGCAACTACGCGCCCGTCGGGGAGCGC
 261  P  R  T  A  D  P  A  V  Q  I  A  G  N  Y  A  P  V  G  E  R

1098 CCGCCCAGCGGCGAGCTCCCGGTGTCCGGCCGCGTCCCGGCGTGCCTCGACGGCGTGTAC
 281  P  P  S  G  E  L  P  V  S  G  R  V  P  A  C  L  D  G  V  Y

1158 GTCCGCAACGGCGCCAACCCGCTCCACGCGCCGCGCGCGGGGCACCACCTCTTCGACGGC
 301  V  R  N  G  A  N  P  L  H  A  P  R  A  G  H  H  L  F  D  G

1218 GACGGCATGCTGCACGCCGTGCGGCTCCGCGGGGGCCGCGCCCAGTCCTACGCGTGCCGG
 321  D  G  M  L  H  A  V  R  L  R  G  G  R  A  Q  S  Y  A  C  R

1278 TTCACGGAGACGGCGCGGCTCCGGCAGGAGCGCGCCCTCGGCAGGGCCGTCTTCCCCAAG
 341  F  T  E  T  A  R  L  R  Q  E  R  A  L  G  R  A  V  F  P  K

1338 GCCATCGGCGAGCTGCACGGCCACTCCGGCGTGGCGCGCCTGCTCCTCTTCGGGGCCCGC
 361  A  I  G  E  L  H  G  H  S  G  V  A  R  L  L  L  F  G  A  R

1398 TCCCTCTGCGGCGTGCTCGACGCGTCGCAGGGGATCGGGGTCGCCAACGCCGGCCTGGTG
```

```
 381  S  L  C  G  V  L  D  A  S  Q  G  I  G  V  A  N  A  G  L  V

1458 TACCATAACAACCGCCTCCTCGCCATGTCGGAGGACGACCTCCCGTACCACGTGCGCGTC
 401  Y  H  N  N  R  L  L  A  M  S  E  D  D  L  P  Y  H  V  R  V

1518 ACCGCCGACGGCGACCTCGAGACCGCCGGGCGCTACGACTTCGGCGGCCAGCTCGACGCC
 421  T  A  D  G  D  L  E  T  A  G  R  Y  D  F  G  G  Q  L  D  A

1578 GCAATGATCGCGCACCCCAAGCTGGACCCGGCCACCGGCGAGCTCTTCGCGCTCAGCTAC
 441  A  M  I  A  H  P  K  L  D  P  A  T  G  E  L  F  A  L  S  Y

1638 AATGTCGTGTCCAGGCCCTTCCTCAAGTACTTCTACTTCACCGCCGACGGCCGCAAGTCC
 461  N  V  V  S  R  P  F  L  K  Y  F  Y  F  T  A  D  G  R  K  S

1698 CCCGACGTCGAGATCCCCGTCGACGCGCCCACCATGATGCACGACTTCGCCGTCACCGAG
 481  P  D  V  E  I  P  V  D  A  P  T  M  M  H  D  F  A  V  T  E

1758 AACCACGCCATCATCCCCGACCAGCAGATCGTGTTCAAGCTCCAGGAGATGGTGCTGGGC
 501  N  H  A  I  I  P  D  Q  Q  I  V  F  K  L  Q  E  M  V  L  G

1818 GGCTCCCCCGTGGTGTACGACAGGAACAAGACCGCGCGGTTCGGTGTGCTGCCGAAGCGC
 521  G  S  P  V  V  Y  D  R  N  K  T  A  R  F  G  V  L  P  K  R

1878 GCCACCGACGCGTCGCGGCTGCAGTGGGTGGACGTCCCCGACTGCTTCTGCTTCCACCTC
 541  A  T  D  A  S  R  L  Q  W  V  D  V  P  D  C  F  C  F  H  L

1938 TGGAACGCGTGGGAGGACGAGGCCACGGGCGAGATCGTGGTGATCGGGTCCTGCATGACC
 561  W  N  A  W  E  D  E  A  T  G  E  I  V  V  I  G  S  C  M  T

1998 CCGGCCGACGCCGTGTTCAACGAGTCCGCCGGCGAGGAGAGCTTCCGCAGCGTGCTGTCG
 581  P  A  D  A  V  F  N  E  S  A  G  E  E  S  F  R  S  V  L  S

2058 GAGATCCGCCTGGACCCGCGCACCGGCACGTCCTCGCGGCGCGCGGTCCTGAGCGACGCC
 601  E  I  R  L  D  P  R  T  G  T  S  S  R  R  A  V  L  S  D  A

2118 GACCAGGTGAACCTGGAGGCCGGCATGGTGAACCGGCAGCTGCTGGGCAGGAAGACGCGC
 621  D  Q  V  N  L  E  A  G  M  V  N  R  Q  L  L  G  R  K  T  R

2178 TACGCCTACCTCGCCATCGCCGAGCCGTGGCCCAAGGTGTCGGGCTTCGCCAAGGTGGAC
 641  Y  A  Y  L  A  I  A  E  P  W  P  K  V  S  G  F  A  K  V  D

2238 CTCGAGGCCGGCACCGTGGAGAAGTTCGTCTACGGCGAGGGCCGGTACGGCGGCGAGCCC
 661  L  E  A  G  T  V  E  K  F  V  Y  G  E  G  R  Y  G  G  E  P

2298 TGCTTCGTGCCGCGCCCGGACGCCCCGGCCGCGGAGGATGACGGGTACGTGCTGTGCTAC
 681  C  F  V  P  R  P  D  A  P  A  A  E  D  D  G  Y  V  L  C  Y

2358 GTGCACGACGAGGGCCGCGGCGCGTCGGAAATGCTCGTCGTCAACGCCCGCGACATGCGG
 701  V  H  D  E  G  R  G  A  S  E  M  L  V  V  N  A  R  D  M  R

2418 GCCGAGGCCGCCGTGAAGCTGCCGAGCCGCGTCCCGTACGGGTTGCACGGCACGTTCATC
 721  A  E  A  A  V  K  L  P  S  R  V  P  Y  G  L  H  G  T  F  I

2478 GCCGGCAAGGAGCTGCAGCGGCAGGCCTAG
 741  A  G  K  E  L  Q  R  Q  A  *
```

10.2.5 ZmNCED9

```
258 ATGTCCCCTATACTAGGTTATTGGAAAATTAAGGGCCTTGTGCAACCCACTCGACTTCTT
  1  M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L

318 TTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAA
 21  L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K

378 TGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT
 41  W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D
```

```
 438 GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAAC
  61  G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N

 498 ATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTG
  81  M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L

 558 GATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTT
 101  D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V

 618 GATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA
 121  D  F  L  S  K  L  P  E  M  L  K  M  F  E  D  R  L  C  H  K

 678 ACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGAT
 141  T  Y  L  N  G  D  H  V  T  H  P  D  F  M  L  Y  D  A  L  D

 738 GTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAA
 161  V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K

 798 AAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCA
 181  K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A

 858 TGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
 201  W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D

 918 CTGGTTCCGCGTGGATCCGCCTCCAACTCCGTCAGGTTCTCGCCGCGCGCCGTCCGCTCC
 221  L  V  P  R  G  S  A  S  N  S  V  R  F  S  P  R  A  V  R  S

 978 GTGCCGCACGAGTGCCTCCAGGCGCCGTTCCACGCCGACCTGCCGGCACCGTCCAAGAAG
 241  V  P  H  E  C  L  Q  A  P  F  H  A  D  L  P  A  P  S  K  K

1038 CCCACCGCCATTGCCGTCCCGAGGCACGCCGCGGCGCCGCGCAAGTCTGGCGGCGGCGGC
 261  P  T  A  I  A  V  P  R  H  A  A  A  P  R  K  S  G  G  G  G

1098 GGCAAGAAGCAGCTCAACCTATTCCAGCGCGCGGCGGCGGCCGCGCTCGACGCGTTCGAG
 281  G  K  K  Q  L  N  L  F  Q  R  A  A  A  A  A  L  D  A  F  E

1158 GAGGGGTTCGTGGCCAACGTCCTCGAGCGGCCCCGCGGGCTGCCCAGCACGGCCGACCCA
 301  E  G  F  V  A  N  V  L  E  R  P  R  G  L  P  S  T  A  D  P

1218 GCCGTGCAGATCGCCGGCAACTTCGCGCCCGTCGGGGAGAGGCTGCCCGTGCATCGGCTC
 321  A  V  Q  I  A  G  N  F  A  P  V  G  E  R  L  P  V  H  R  L

1278 CCCGTCTCCGGCCGCATCCCGCCCTTCATCAGCGGCGTCTACGCGCGCAACGGCGCCAAC
 341  P  V  S  G  R  I  P  P  F  I  S  G  V  Y  A  R  N  G  A  N

1338 CCCTGCTTCGACCCCGTCGCCGGGCACCACCTCTTCGACGGCGACGGCATGGTGCACGCG
 361  P  C  F  D  P  V  A  G  H  H  L  F  D  G  D  G  M  V  H  A

1398 CTGCGGATACGCGACGGCGTCGCCGAGTCCTACGCCTGCCGCTTCACCGAGACCGCGCGC
 381  L  R  I  R  D  G  V  A  E  S  Y  A  C  R  F  T  E  T  A  R

1458 CTGACCCAGGAGCGCGCGGTCGGCCGCCCCGTCTTCCCCAAGGCCATCGGCGAGCTGCAC
 401  L  T  Q  E  R  A  V  G  R  P  V  F  P  K  A  I  G  E  L  H

1518 GGCCACTCCGGGATCGCGCGCCTCGCCCTGTTCTACGCGCGCGCCGCGTGCGGCCTCGTC
 421  G  H  S  G  I  A  R  L  A  L  F  Y  A  R  A  A  C  G  L  V

1578 GACCCCTCGGCCGGCACCGGCGTCGCCAACGCCGGACTCGTCTACTTCAACGGCCGCCTC
 441  D  P  S  A  G  T  G  V  A  N  A  G  L  V  Y  F  N  G  R  L

1638 CTCGCCATGTCCGAGGACGACCTGCCGTACCACGTCCGCGTCGCGGACGACGGCGACCTC
 461  L  A  M  S  E  D  D  L  P  Y  H  V  R  V  A  D  D  G  D  L

1698 GAGACCGTCGGCCGCTACGACTTCGACGGCCAGCTCGGCTGCCCCATGATCGCGCACCCC
 481  E  T  V  G  R  Y  D  F  D  G  Q  L  G  C  P  M  I  A  H  P

1758 AAGCTGGACCCGGCCACCGGGGAGCTGCACGCGCTCAGCTACGAGGTCGTCAGGAGGCCC
```

```
 501  K  L  D  P  A  T  G  E  L  H  A  L  S  Y  E  V  V  R  R  P

1818 TACCTCAAGTACTTCTACTTCAGGCCCGACGGCACCAAGTCCGACGACGTGGAGATCCCG
 521  Y  L  K  Y  F  Y  F  R  P  D  G  T  K  S  D  D  V  E  I  P

1878 CTGGCCCAGCCCACCATGATCCACGACTTCGCCATCACCGAGAACTTGGTCGTGGTGCCC
 541  L  A  Q  P  T  M  I  H  D  F  A  I  T  E  N  L  V  V  V  P

1938 GACCACCAGGTGGTGTTCAAGCTGCAGGAGATGCTGCGCGGCGGGTCGCCCGTGGTGCTG
 561  D  H  Q  V  V  F  K  L  Q  E  M  L  R  G  G  S  P  V  V  L

1998 GACAGGGAGAAGACGTCGCGCTTCGGCGTCCTCCCGAAGCGCGCCGCGGACGCGTCGGAG
 581  D  R  E  K  T  S  R  F  G  V  L  P  K  R  A  A  D  A  S  E

2058 ATGGCGTGGGTGGACGTGCCGGACTGCTTCTGCTTCCACCTGTGGAACGCGTGGGAGGAC
 601  M  A  W  V  D  V  P  D  C  F  C  F  H  L  W  N  A  W  E  D

2118 GAGGCGACGGGCGAGGTGGTGGTGATCGGCTCCTGCATGACCCCCGCCGACTCCATCTTC
 621  E  A  T  G  E  V  V  V  I  G  S  C  M  T  P  A  D  S  I  F

2178 AACGAGTCGGACGTGCGGCTCGAGAGCGTGCTGACGGAGATCCGGCTGGACGCGCGCACG
 641  N  E  S  D  V  R  L  E  S  V  L  T  E  I  R  L  D  A  R  T

2238 GGCCGGTCGACGCGCCGCGCCGTCCTGCCGCCGTCGCGGCAGGTGAACCTGGAGGTGGGC
 661  G  R  S  T  R  R  A  V  L  P  P  S  R  Q  V  N  L  E  V  G

2298 ATGGTGAACCGCAACCTCCTGGGGCGCAGGACGCGGTACGCGTACCTCGCGGTGGCCGAG
 681  M  V  N  R  N  L  L  G  R  R  T  R  Y  A  Y  L  A  V  A  E

2358 CCGTGGCCCAAGGTCTCGGGCTTCGCCAAGGTGGACCTGGCCACGGGCGAGCTCGCCAGG
 701  P  W  P  K  V  S  G  F  A  K  V  D  L  A  T  G  E  L  A  R

2418 TTCGAGTACGGCGAGGGCCGGTTCGGCGGCGAGCCCTGCTTCGTGCCCATGGACCCCGCC
 721  F  E  Y  G  E  G  R  F  G  G  E  P  C  F  V  P  M  D  P  A

2478 GCGGCCCACCCGCGCGGCGAGGACGACGGGCACGTGCTCGCCTTCGTCCACGACGAGCGC
 741  A  A  H  P  R  G  E  D  D  G  H  V  L  A  F  V  H  D  E  R

2538 GCCGGCACGTCCGAGCTCCTGGTGGTCAATGCCGCCGACATGCGGCTGGAGGCCACCGTC
 761  A  G  T  S  E  L  L  V  V  N  A  A  D  M  R  L  E  A  T  V

2598 CGGCTCCCGTCCCGCGTGCCCTTCGGCTTCCACGGCACCTTCATCACGGGCGCGGAGCTC
 781  R  L  P  S  R  V  P  F  G  F  H  G  T  F  I  T  G  A  E  L

2658 GAGGCCCAGGCCTGA
 801  E  A  Q  A  *
```

10.3 *Arabidopsis thaliana* ABA2 Optimised Sequence

Accession number of native *A. thaliana* ABA2 sequence: BT003412. Optimised sequence contains no introns. His tag is shown in black. Main ABA2 sequence is shown in red.

```
297 ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
  1  M  R  G  S  H  H  H  H  H  H  G  M  A  S  M  T  G  G  Q  Q

357 ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTCACCATGTCAACGAAT
 21  M  G  R  D  L  Y  D  D  D  D  K  D  H  P  F  T  M  S  T  N

417 ACCGAATCCTCTTCTTACTCATCTCTGCCGAGCCAACGCCTGCTGGGCAAAGTCGCTCTG
 41  T  E  S  S  S  Y  S  S  L  P  S  Q  R  L  L  G  K  V  A  L

477 ATTACGGGTGGTGCGACCGGCATTGGTGAATCTATCGTTCGTCTGTTCCATAAACACGGC
 61  I  T  G  G  A  T  G  I  G  E  S  I  V  R  L  F  H  K  H  G
```

```
 537 GCTAAAGTGTGCATTGTTGATCTGCAGGATGACCTGGGCGGTGAAGTGTGTAAATCCCTG
  81  A  K  V  C  I  V  D  L  Q  D  D  L  G  G  E  V  C  K  S  L

 597 CTGCGTGGCGAATCAAAAGAAACCGCATTTTTCATTCATGGTGATGTCCGCGTGGAAGAT
 101  L  R  G  E  S  K  E  T  A  F  F  I  H  G  D  V  R  V  E  D

 657 GACATCTCAAACGCCGTGGATTTCGCAGTTAAAAATTTTGGTACGCTGGACATTCTGATC
 121  D  I  S  N  A  V  D  F  A  V  K  N  F  G  T  L  D  I  L  I

 717 AACAATGCTGGTCTGTGCGGTGCACCGTGTCCGGATATCCGTAACTATAGCCTGTCTGAA
 141  N  N  A  G  L  C  G  A  P  C  P  D  I  R  N  Y  S  L  S  E

 777 TTCGAAATGACCTTTGACGTTAATGTCAAAGGCGCCTTTCTGTCCATGAAACACGCGGCC
 161  F  E  M  T  F  D  V  N  V  K  G  A  F  L  S  M  K  H  A  A

 837 CGCGTTATGATTCCGGAGAAAAAAGGTAGTATCGTCTCCCTGTGCTCAGTGGGTGGTGTG
 181  R  V  M  I  P  E  K  K  G  S  I  V  S  L  C  S  V  G  G  V

 897 GTTGGTGGTGTTGGTCCGCATTCTTATGTCGGTAGTAAACACGCAGTGCTGGGTCTGACG
 201  V  G  G  V  G  P  H  S  Y  V  G  S  K  H  A  V  L  G  L  T

 957 CGTTCGGTTGCAGCTGAACTGGGCCAACATGGTATTCGCGTCAACTGTGTGAGCCCGTAC
 221  R  S  V  A  A  E  L  G  Q  H  G  I  R  V  N  C  V  S  P  Y

1017 GCTGTGGCGACCAAACTGGCCCTGGCACACCTGCCGGAAGAAGAACGTACGGAAGATGCC
 241  A  V  A  T  K  L  A  L  A  H  L  P  E  E  E  R  T  E  D  A

1077 TTTGTGGGCTTCCGCAACTTTGCGGCCGCAAACGCAAATCTGAAAGGTGTTGAACTGACG
 261  F  V  G  F  R  N  F  A  A  A  N  A  N  L  K  G  V  E  L  T

1137 GTTGATGACGTCGCTAATGCGGTCCTGTTCCTGGCCTCGGATGACAGCCGCTACATCAGC
 281  V  D  D  V  A  N  A  V  L  F  L  A  S  D  D  S  R  Y  I  S

1197 GGTGATAACCTGATGATCGACGGCGGCTTCACCTGCACCAACCATTCGTTCAAAGTGTTT
 301  G  D  N  L  M  I  D  G  G  F  T  C  T  N  H  S  F  K  V  F

1257 CGCTGA
 321  R  *
```

10.4 *Oryza Sativa* D27 Optimised Sequence

Accession number of native sequence: FJ641055. Optimised sequence contains no introns. His tag is shown in black. Main ABA2 sequence is shown in red. For GST tagged construct, GST tag is the same as shown for *Z. mays* NCED sequences. Truncated protein lacks the first 120 base pairs of the red sequence.

```
   1 ATGCGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAA
   1  M  R  G  S  H  H  H  H  H  H  G  M  A  S  M  T  G  G  Q  Q

  61 ATGGGTCGGGATCTGTACGACGATGACGATAAGGATCATCCCTTCACCATGGAAACGACG
  21  M  G  R  D  L  Y  D  D  D  D  K  D  H  P  F  T  M  E  T  T

 121 ACCCTGGTTCTGCTGCTGCCGCATGGTGGTGCTGGCGGTGTGCGTCCGGCTGCTGCTGCA
  41  T  L  V  L  L  L  P  H  G  G  A  G  G  V  R  P  A  A  A  A

 181 ACGGCGAAGCGTAGCTATGTTATGCGTCGCTGCTGTTCTACCGTCCGTGCAGTGATGGCA
  61  T  A  K  R  S  Y  V  M  R  R  C  C  S  T  V  R  A  V  M  A

 241 CGCCCGCAGGAAGCCCCGGCAAGTGCTCCGGCGAAAAAGACCGAAACGGCGGCCATGATG
  81  R  P  Q  E  A  P  A  S  A  P  A  K  K  T  E  T  A  A  M  M

 301 TCCACCGTTCAAACCGAAACGGCAGCTGCGCCGCCGGCTACGGTCTATCGTGATAGTTGG
 101  S  T  V  Q  T  E  T  A  A  A  P  P  A  T  V  Y  R  D  S  W
```

```
361 TTTGACAAACTGGCCATTGGTTACCTGTCACGCAACCTGCAGGAAGCATCGGGCCTGAAA
121  F  D  K  L  A  I  G  Y  L  S  R  N  L  Q  E  A  S  G  L  K

421 AATGAAAAGGATGGTTACGAAAGCCTGATTGACGCCGCACTGGCAATTAGTCGTATCTTT
141  N  E  K  D  G  Y  E  S  L  I  D  A  A  L  A  I  S  R  I  F

481 TCCCTGGATAAACAGTCAGAAATCGTGACCCAAGCTCTGGAACGCGCGCTGCCGTCGTAT
161  S  L  D  K  Q  S  E  I  V  T  Q  A  L  E  R  A  L  P  S  Y

541 ATTCTGACCATGATCAAGGTTATGATGCCGCCGAGCCGTTTCTCTCGCGAATACTTTGCT
181  I  L  T  M  I  K  V  M  M  P  P  S  R  F  S  R  E  Y  F  A

601 GCGTTCACCACGATTTTCTTTCCGTGGCTGGTTGGCCCGTGCGAAGTCATGGAAAGCGAA
201  A  F  T  T  I  F  F  P  W  L  V  G  P  C  E  V  M  E  S  E

661 GTGGAAGGTCGTAAAGAAAAGAACGTGGTTTATATCCCGAAATGTCGCTTTCTGGAATCT
221  V  E  G  R  K  E  K  N  V  V  Y  I  P  K  C  R  F  L  E  S

721 ACCAACTGCGTCGGCATGTGTACGAATCTGTGCAAAATTCCGTGTCAGAAGTTCATCCAA
241  T  N  C  V  G  M  C  T  N  L  C  K  I  P  C  Q  K  F  I  Q

781 GATAGCCTGGGTATGAAAGTGTACATGAGCCCGAATTTTGAAGACATGTCTTGCGAAATG
261  D  S  L  G  M  K  V  Y  M  S  P  N  F  E  D  M  S  C  E  M

841 ATCTTCGGTCAGCAACCGCCGGAAGATGACCCGGCCCTGAAACAGCCGTGTTTCCGTACC
281  I  F  G  Q  Q  P  P  E  D  D  P  A  L  K  Q  P  C  F  R  T

901 AAGTGCGTTGCGAAGCAGAATCATGGCGTGAACTGTAGCATTTGA
301  K  C  V  A  K  Q  N  H  G  V  N  C  S  I  *
```

10.5 Sequence of *Arabidopsis thaliana* CCD7

Accession number of native sequence: NM_130064. Sequence of CCD7.1 and CCD7.2 differs from that of NM_130064 from base pair 1793. Sequence of CCD7 used during this investigation is shown below. GST-tag is shown in black. Main CCD7 sequence is shown in red. Truncated protein lacks the first 93 base pairs of the red sequence.

```
  1  M  S  P  I  L  G  Y  W  K  I  K  G  L  V  Q  P  T  R  L  L

318 TTGGAATATCTTGAAGAAAAATATGAAGAGCATTTGTATGAGCGCGATGAAGGTGATAAA
 21  L  E  Y  L  E  E  K  Y  E  E  H  L  Y  E  R  D  E  G  D  K

378 TGGCGAAACAAAAAGTTTGAATTGGGTTTGGAGTTTCCCAATCTTCCTTATTATATTGAT
 41  W  R  N  K  K  F  E  L  G  L  E  F  P  N  L  P  Y  Y  I  D

438 GGTGATGTTAAATTAACACAGTCTATGGCCATCATACGTTATATAGCTGACAAGCACAAC
 61  G  D  V  K  L  T  Q  S  M  A  I  I  R  Y  I  A  D  K  H  N

498 ATGTTGGGTGGTTGTCCAAAAGAGCGTGCAGAGATTTCAATGCTTGAAGGAGCGGTTTTG
 81  M  L  G  G  C  P  K  E  R  A  E  I  S  M  L  E  G  A  V  L

558 GATATTAGATACGGTGTTTCGAGAATTGCATATAGTAAAGACTTTGAAACTCTCAAAGTT
101  D  I  R  Y  G  V  S  R  I  A  Y  S  K  D  F  E  T  L  K  V

618 GATTTTCTTAGCAAGCTACCTGAAATGCTGAAAATGTTCGAAGATCGTTTATGTCATAAA
121  D  F  L  S  K  L  P  E  M  L  K  M  F  E  D  R  L  C  H  K

678 ACATATTTAAATGGTGATCATGTAACCCATCCTGACTTCATGTTGTATGACGCTCTTGAT
141  T  Y  L  N  G  D  H  V  T  H  P  D  F  M  L  Y  D  A  L  D

738 GTTGTTTTATACATGGACCCAATGTGCCTGGATGCGTTCCCAAAATTAGTTTGTTTTAAA
161  V  V  L  Y  M  D  P  M  C  L  D  A  F  P  K  L  V  C  F  K
```

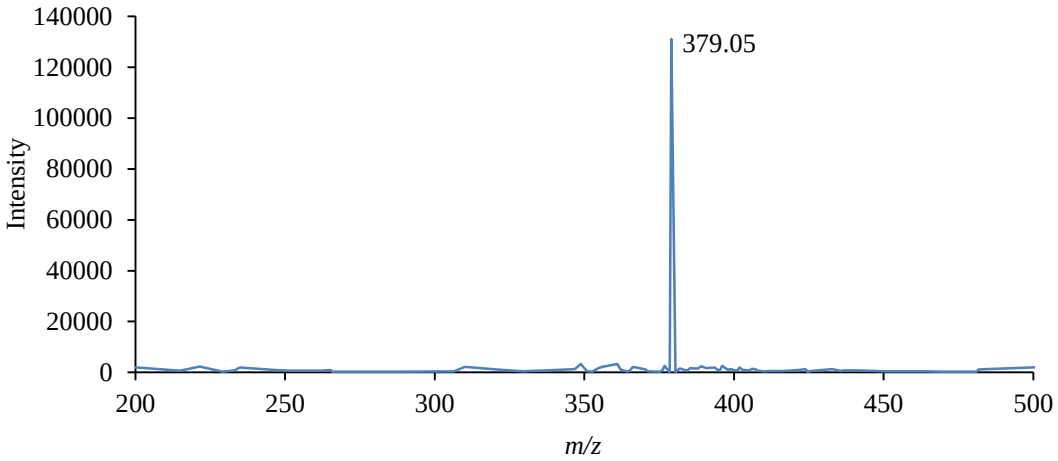
```
 798 AAACGTATTGAAGCTATCCCACAAATTGATAAGTACTTGAAATCCAGCAAGTATATAGCA
 181  K  R  I  E  A  I  P  Q  I  D  K  Y  L  K  S  S  K  Y  I  A

 858 TGGCCTTTGCAGGGCTGGCAAGCCACGTTTGGTGGTGGCGACCATCCTCCAAAATCGGAT
 201  W  P  L  Q  G  W  Q  A  T  F  G  G  G  D  H  P  P  K  S  D

 918 CTGGTTCCGCGTGGATCCGGATTAATGTCTCTCCCTATCCCGCCGAAATTTCTTCCACCG
 221  L  V  P  R  G  S  G  L  M  S  L  P  I  P  P  K  F  L  P  P

 978 CTAAAATCTCCACCGATTCATCATCACCAAACTCCGCCACCGCTTGCACCTCCACGAGCC
 241  L  K  S  P  P  I  H  H  H  Q  T  P  P  P  L  A  P  P  R  A

1038 GCAATATCAATATCTATACCAGACACCGGTTTAGGACGTACCGGTACCATCCTCGACGAG
 261  A  I  S  I  S  I  P  D  T  G  L  G  R  T  G  T  I  L  D  E

1098 TCCACGTCTTCAGCTTTCCGTGATTACCAATCTTTATTCGTGTCACAACGTTCCGAGACT
 281  S  T  S  S  A  F  R  D  Y  Q  S  L  F  V  S  Q  R  S  E  T

1158 ATCGAACCGGTCGTAATTAAACCAATCGAAGGTTCAATACCGGTTAACTTCCCTTCCGGT
 301  I  E  P  V  V  I  K  P  I  E  G  S  I  P  V  N  F  P  S  G

1218 ACATACTACTTAGCCGGTCCAGGACTATTTACTGACGACCATGGCTCAACGGTTCATCCT
 321  T  Y  Y  L  A  G  P  G  L  F  T  D  D  H  G  S  T  V  H  P

1278 TTAGACGGTCACGGTTATCTCCGTGCGTTTCACATCGACGGTAACAAACGGAAAGCCACT
 341  L  D  G  H  G  Y  L  R  A  F  H  I  D  G  N  K  R  K  A  T

1338 TTCACGGCGAAGTACGTTAAAACGGAAGCTAAAAAAGAAGAGCACGATCCTGTAACTGAC
 361  F  T  A  K  Y  V  K  T  E  A  K  K  E  E  H  D  P  V  T  D

1398 ACGTGGCGGTTCACTCATAGAGGTCCTTTCTCGGTGTTGAAAGGAGGGAAGAGATTTGGA
 381  T  W  R  F  T  H  R  G  P  F  S  V  L  K  G  G  K  R  F  G

1458 AACACGAAAGTGATGAAAAACGTGGCTAATACTAGCGTTTTGAAATGGGCTGGGCGATTG
 401  N  T  K  V  M  K  N  V  A  N  T  S  V  L  K  W  A  G  R  L

1518 CTTTGTTTATGGGAAGGTGGTGAACCGTACGAGATTGAATCTGGATCGTTGGATACCGTC
 421  L  C  L  W  E  G  G  E  P  Y  E  I  E  S  G  S  L  D  T  V

1578 GGAAGATTTAACGTCGAGAACAACGGTTGTGAATCTTGTGATGATGATGATTCTTCCGAC
 441  G  R  F  N  V  E  N  N  G  C  E  S  C  D  D  D  D  S  S  D

1638 AGAGATTTATCTGGTCATGATATATGGGACACAGCCGCAGATTTGTTGAAACCCATACTT
 461  R  D  L  S  G  H  D  I  W  D  T  A  A  D  L  L  K  P  I  L

1698 CAAGGTGTATTTAAGATGCCACCGAAACGGTTCTTGTCACATTACAAAGTCGACGGTCGA
 481  Q  G  V  F  K  M  P  P  K  R  F  L  S  H  Y  K  V  D  G  R

1758 AGAAAAAGACTTTTAACGGTCACTTGCAACGCTGAAGATATGCTTTTACCTCGAAGCAAC
 501  R  K  R  L  L  T  V  T  C  N  A  E  D  M  L  L  P  R  S  N

1818 TTCACATTTTGTGAGTATGATTCGGAATTCAAGTTGATACAAACGAAAGAATTCAAGATC
 521  F  T  F  C  E  Y  D  S  E  F  K  L  I  Q  T  K  E  F  K  I

1878 GATGATCATATGATGATTCATGATTGGGCATTCACGGATACTCACTACATACTCTTTGCC
 541  D  D  H  M  M  I  H  D  W  A  F  T  D  T  H  Y  I  L  F  A

1938 AACCGAGTCAAGCTTAATCCAATAGGTTCCATAGCGGCTATGTGCGGAATGTCACCAATG
 561  N  R  V  K  L  N  P  I  G  S  I  A  A  M  C  G  M  S  P  M

1998 GTATCAGCGTTATCGTTAAACCCAAGCAACGAGAGTTCTCCGATTTATATTCTCCCTAGG
 581  V  S  A  L  S  L  N  P  S  N  E  S  S  P  I  Y  I  L  P  R

2058 TTTTCTGATAAATATTCTAGGGGAGGTCGAGACTGGAGAGTTCCTGTCGAAGTGTCTTCT
 601  F  S  D  K  Y  S  R  G  G  R  D  W  R  V  P  V  E  V  S  S

2118 CAATTATGGCTAATACACTCCGGAAACGCTTATGAGACTAGAGAGGATAACGGCGATTTA
 621  Q  L  W  L  I  H  S  G  N  A  Y  E  T  R  E  D  N  G  D  L

2178 AAGATTCAGATACAAGCTTCCGCTTGTTCTTACCGATGGTTCGATTTCCAGAAAATGTTT
 641  K  I  Q  I  Q  A  S  A  C  S  Y  R  W  F  D  F  Q  K  M  F
```

```
2238 GGCTATGATTGGCAAAGCAACAAGCTGGATCCTTCTGTTATGAATCTAAACCGTGGCGAC
 661  G  Y  D  W  Q  S  N  K  L  D  P  S  V  M  N  L  N  R  G  D

2298 GACAAACTACTCCCTCATCTAGTTAAGGTGTCTATGACTTTGGACTCTACCGGAAACTGC
 681  D  K  L  L  P  H  L  V  K  V  S  M  T  L  D  S  T  G  N  C

2358 AATAGTTGTGATGTAGAGCCTCTAAACGGGTGGAACAAGCCGTCAGATTTTCCGGTTATA
 701  N  S  C  D  V  E  P  L  N  G  W  N  K  P  S  D  F  P  V  I

2418 AACTCATCCTGGTCCGGAAAAAAGAACAAGTACATGTACTCTGCTGCCTCGTCGGGAACT
 721  N  S  S  W  S  G  K  K  N  K  Y  M  Y  S  A  A  S  S  G  T

2478 CGAAGTGAACTTCCCCATTTTCCATTCGACATGGTCGTGAAATTTGACTTAGACTCAAAC
 741  R  S  E  L  P  H  F  P  F  D  M  V  V  K  F  D  L  D  S  N

2538 CTCGTCCGTACTTGGTCTACCGGAGCTAGAAGATTCGTTGGTGAGCCCATGTTTGTCCCA
 761  L  V  R  T  W  S  T  G  A  R  R  F  V  G  E  P  M  F  V  P

2598 AAAAACTCTGTTGAAGAAGGAGAAGAAGAGGACGATGGTTACATTGTCGTGGTCGAGTAT
 781  K  N  S  V  E  E  G  E  E  E  D  D  G  Y  I  V  V  V  E  Y

2658 GCGGTTTCGGTGGAGAGATGTTACCTAGTGATTTTGGATGCTAAGAAGATCGGTGAATCC
 801  A  V  S  V  E  R  C  Y  L  V  I  L  D  A  K  K  I  G  E  S

2718 GATGCGGTCGTGTCGAGGTTAGAGGTTCCGAGGAATTTGACGTTTCCGATGGGTTTTCAT
 821  D  A  V  V  S  R  L  E  V  P  R  N  L  T  F  P  M  G  F  H

2778 GGTTTATGGGCTAGCGACTGA
 841  G  L  W  A  S  D  *
```

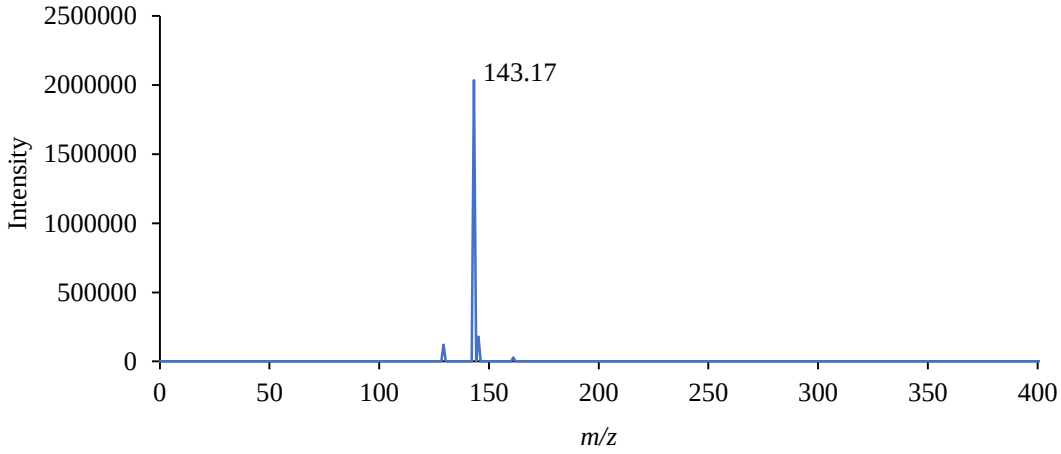
10.6 Low Resolution Mass Spectrometry Spectrum for CCD8 Cleavage Reaction Breakdown Products

Spectra were obtained via direct infusion using a Bruker HTC-Ultra electrospray ionisation mass spectrometer and trapped using an ion trap.

10.6.1 Spectra for all-trans-β-apo-10'-carotenol



10.6.2 Spectra for (3E)-4-(2,6,6-Trimethylcyclohex-1-en-1-yl)2-methyl-but-3-en-1-ol



10.6.3 Spectra for 4-methylhept-2,4-en-1,7-diol